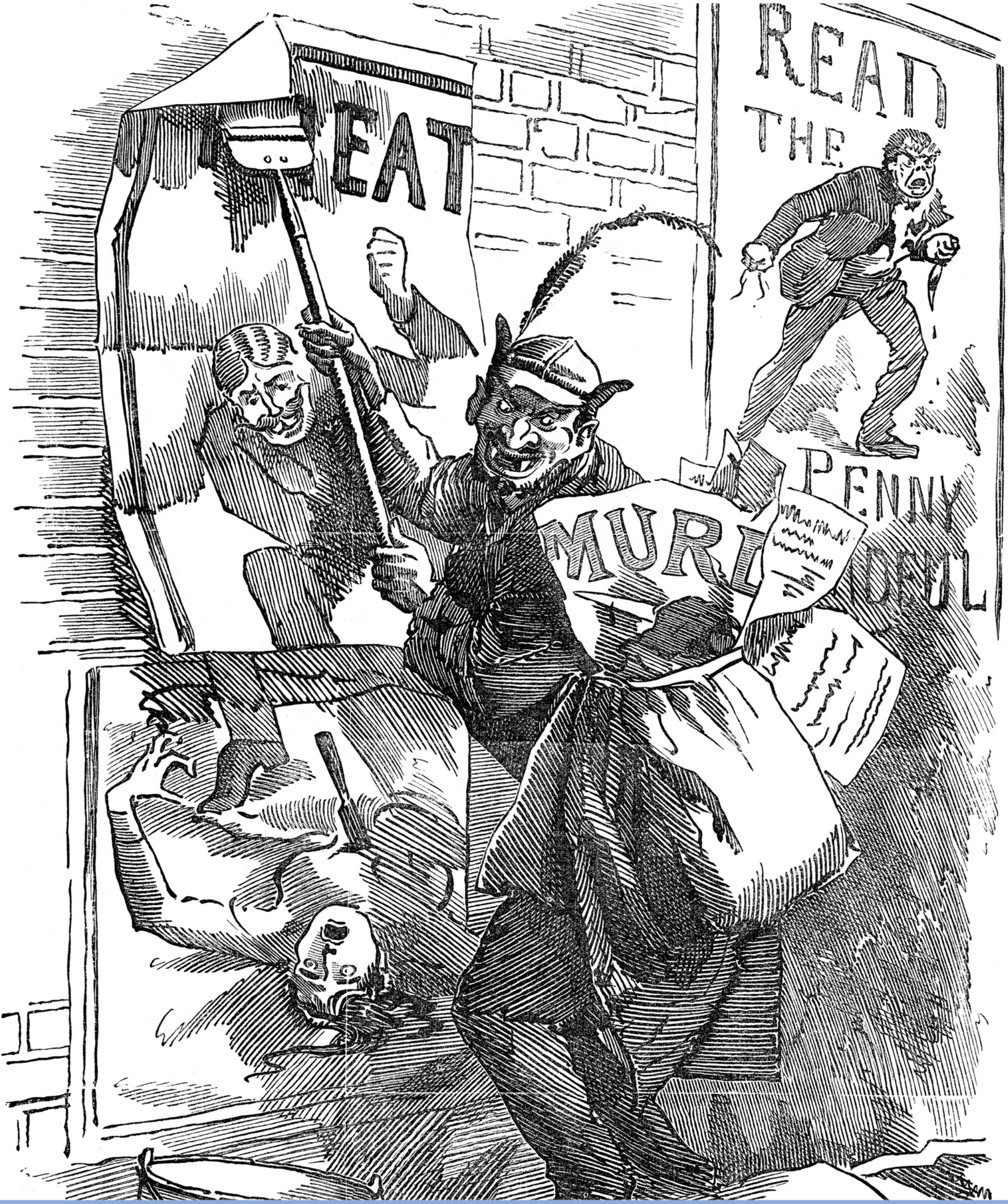

FICTIONAL IMMORALITY AND IMMORAL FICTION

GARRY YOUNG

Copyright 2021. Lexington Books. All rights reserved. May not be reproduced in any form without permission from the publisher, except fair uses permitted under U.S. or applicable copyright law.

EBSCO Publishing : eBook Collection (EBSCOhost) - printed on 2/12/2023 3:24 AM via
AN: 2729433 ; Garry Young.; Fictional Immorality and Immoral Fiction
Account: ns335141

Fictional Immorality and Immoral Fiction

Fictional Immorality and Immoral Fiction

Garry Young

LEXINGTON BOOKS
Lanham • Boulder • New York • London

Published by Lexington Books
An imprint of The Rowman & Littlefield Publishing Group, Inc.
4501 Forbes Boulevard, Suite 200, Lanham, Maryland 20706
www.rowman.com

6 Tinworth Street, London SE11 5AL, United Kingdom

Copyright © 2021 by The Rowman & Littlefield Publishing Group, Inc.

All rights reserved. No part of this book may be reproduced in any form or by any electronic or mechanical means, including information storage and retrieval systems, without written permission from the publisher, except by a reviewer who may quote passages in a review.

British Library Cataloguing in Publication Information Available

Library of Congress Control Number: 2020946695

ISBN 978-1-7936-3919-6 (cloth : alk. paper)
ISBN 978-1-7936-3920-2 (electronic)

♾™ The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI/NISO Z39.48-1992.

Contents

1 Introduction: Fictional Immorality and Immoral Fiction 1

2 The Metaphysics of Fictional Immorality 19

3 The Content of Fictional Immorality 35

4 The Meaning of Fictional Immorality 55

5 The Harm of Fictional Immorality 77

6 Is There Evidence of Harm? 99

7 Enjoying Fictional Immorality 115

8 Resisting Fictional Immorality 133

9 Poor Taste and Fictional Immorality 147

10 Historical Fiction and Fictional Immorality 159

11 A New Kind of Fiction 177

12 Fantasy and Fictional Immorality 187

13 "It's Not Immoral, but It Is in Poor Taste" 207

14 Immoral Fiction and Censorship 225

References 241

Index 263

About the Author 269

Chapter 1

Introduction

Fictional Immorality and Immoral Fiction

FICTIONAL IMMORALITY

Fictional immorality can take many forms. A common example would be the murder of a character in a film or a television program, a work of literature or a play, or perhaps even a comic book, song lyric, or video game. Just as easily, the immorality depicted could be theft—an outlandish bank robbery, perhaps—or sexual assault or torture, or any number of other immoral actions. The point is: for x to be an example of *fictional* immorality (or "fictional-x_i"), first and foremost, it must *depict* something that, if carried out for real, would be immoral (likely illegal). Moreover, as part of a "strict" definition of fictional immorality, fictional-x_i must involve a fictitious event with only fictitious characters (i.e., it must not depict *actual* individuals or events). The murder of Nancy by Bill Sikes in Charles Dickens' novel *Oliver Twist* satisfies this strict definition, whereas the assassination of Julius Caesar in Shakespeare's eponymously titled play does not.

As we progress, it is my intention to loosen this definition by allowing depictions of fictional immorality to be situated within historical contexts (say, a fictitious event that takes place during World War II) or to involve historical figures (e.g., a fictitious plot to murder a famous composer). Initially restricting fictional-x_i to *fictitious* characters engaged in *fictitious* acts of immorality is, however, an important first step in the scrutiny of fictional immorality because it allows conclusions to be drawn about the moral status of fictional immorality in its purest form, or what Nolan and Sandgren (2014), following Parsons (1980), call *native fiction*, even though such fiction may not be completely removed from reality. The murder of Nancy by Bill Sikes is, after all, set in London. The subsequent and systematic "contamination" of "pure fiction" with nonfiction-related properties, or what Nolan and Sandgren

(2014)—again, following Parsons (1980)—refer to as *immigrant fiction*, is therefore designed to reveal the impact of these additional properties on certain, yet to be discussed, morally pertinent matters.

Depicting Fiction

Fictional-x_i *depicts* an immoral action (x_i) that is not directed toward nor does it include in any significant way, the depiction of an actual object, animal, person, or event, or otherwise allude to any of these.[1] A cartoon sketch of the crew of the Enola Gay releasing the atomic bomb over Hiroshima does not constitute fictional-x_i, whereas a limerick about the Wizard of Oz bludgeoning to death the Tin Man in front of a shocked Dorothy does. Likewise, a caricature of Donald Trump cannibalizing the corpses of Latin American illegal immigrants, even though it is understood to be a fiction (*qua* the product of the artist's imagination), fails to satisfy the strict definition I am employing here. The same can be said of a painting of Jesus of Nazareth crucified, as well as S's fantasy about his (nonexistent) love affair with his next-door neighbor's wife.

What should be apparent from these examples is that I intend to adopt a fairly broad definition of "depiction." It is my contention, however, that such inclusivity will not hamper the argument to come. That said, in order to avoid incorporating an overly liberal notion of "depiction" within my definition of fictional-x_i, I shall treat private fantasies—such as S's private fantasy about having an extramarital affair with the cartoon character Marge Simpson—as a separate category of depiction. This is because while it is true that S's fantasy involves fictional characters (in this case, Marge Simpson and quite possibly her betrayed husband, Homer), it is not a fiction that is depicted in an *overt* manner (i.e., in the form of a drawing, painting, story, diary entry, or song, or any other physical manifestation, including behavior). In contrast, if S were to enact the fantasy while playing the central (fictional) character in a self-penned play about sexual fantasy, then it would adhere, more typically, to the notion of depiction I intend to employ throughout the majority of this book.

Engaging with Fiction

What is it to *engage* with fiction? There are a number of ways one might answer this question, each equally valid. One might examine the author/creator's writing or animation style, for example, or deconstruct their lyrical or joke content and structure, or perhaps admire the way in which some new technology has been applied to create *this* fiction in *this* way (e.g., CGI), rather than some more traditional manner (e.g., frame-by-frame hand-drawn animation). None of these modes of engagement is of interest, however, because none of them captures fully what it is to engage with the fiction as a work of *fiction*, as opposed to some technical exercise in digital animation or the use of narration in the first-person plural, and so on. That is, none *requires* that one treat the product as a fiction, rather than as a way to understand how one creates the product, whatever the product happens to be. The first step on the road to engaging with fiction *qua* fiction, then, is the use of imagination: for, as Stock (2016) notes, fiction prescribes imagining.

Following Gaut (2003), minimally, the act of imagining something—say, that one is conversing with a grinning Cheshire cat—is to entertain the propositional or nonpropositional content without committing to its truth or falsity (without, that is, alethic commitment). Engaging one's imagination in this sense, however, while necessary, is not sufficient for the sort of imagining I have in mind. To illustrate: When imagining the outcome of a decision I have not yet made—for example, what might happen if I decide to call off my wedding—while it is true that I am entertaining a fiction of sorts, *qua* a hypothetical event, without committing to its truth or falsity, this type of fiction is hardly representative of the fiction found in (*inter alia*) works of literature, film and television dramas, plays, video games, poetry, and songs. To engage with these kinds of fiction requires that one *fictively imagine x* (Cooke, 2014): that is, engage with the depiction without making nontrivial inferences about the beliefs of the author on the basis of depiction *x*.

From my act of *imagining* calling off the wedding, and my contemplation of how, hypothetically, this decision would play out, at the very least, I might infer that I had "cold feet" and therefore I *believed* I was making a mistake by getting married, or at least this possibility was something I was examining. However, if I were to read a novel about a man who, say, championed the bachelor lifestyle, then in order to engage with this description as a fiction *qua* fictively imagining this man, not only would I *not* commit to the truth or falsity of the fictional character and his championing of bachelorhood; more than this, I would *not* make inferences about the author of the fiction from what is depicted: that the *author* believes in the merits of a bachelor lifestyle over a committed relationship.

Following Stock (2016), to fictively imagine that *p* is to think (consciously) that *p* in a way that represents *p*'s occurrence without committing to the truth of *p*. Fictively imagining that *p*—say, that I am reading at my desk—requires either that I do not believe that *p* (for I cannot imagine *only* what I also believe) or that I believe that *p* but I am disposed to connect this belief, via inference, to some further propositional content that I imagine but do not believe: that I am living on the moon and I am therefore reading at my desk at this location.

Let us convert these points into more formal conditions. S is able to fictively imagine that *p if* and *only if* conditions (i) and (ii) are satisfied:

(i) S lacks alethic commitment to *p* or, where S simultaneously believes that *p*, the sum of S's thoughts to which she is disposed to be inferentially connected is different in terms of what she imagines and what she believes, resulting in an overall deficit in alethic commitment.
(ii) Where S is not the creator of the depiction of *p* (that she is fictively imagining), she does not infer from the depiction that the creator is alethically committed to *p*.

In accordance with conditions (i) and (ii), when reading about or viewing the exploits of private detective Cormoran Strike (the lead character in a fictional detective series written by Robert Galbraith), I will fictively imagine these exploits only insofar as I lack alethic commitment to them *and* fail to infer the same alethic commitment on the part of the author. Equally, I will fictively imagine that I am reading about the exploits of Cormoran Strike *while sat at my desk*, even if I *believe* that I am sitting at my desk, only if I also imagine doing this while commuting to the moon (for example) in the latest hyper-speed space shuttle: something I do not believe is happening but which constitutes a thought that is inferentially connected to my belief about being sat at my desk during my episode of fictive imagining.

When I engage with fiction by fictively imagining Cormoran Strike unraveling some criminal plot, or the exploits of action hero Deadpool, complete with his sardonic humor, or Ripley's claustrophobic struggle with the alien on board the Nostromo, I take on the role of an avid consumer. There are other ways to engage with fiction *qua* fiction, of course—*inter alia*, as a producer or a distributor—each of which is not immune to moral scrutiny. My interest, however, is *mostly* in the morality of engaging with fictional-x_i as a consumer. I say "mostly" because such interest still requires that I consider the act of creating the fiction, especially when discussing the meaning I glean from the fiction as a consumer (i.e., how does it align with the creator's intended meaning, if known?). In short, then, while I do not deny that producing, storing, distributing, and selling fictional immorality are forms of engagement, nor that they raise their own morally pertinent issues, they are not of concern here.

Different Media

With the exception of private fantasy, which I will exclude from initial consideration and discuss separately (for reasons already mentioned), it is my intention to treat different media—*inter alia* films, TV, literature, video games—as nominally homogenous unless otherwise stated. Different media will therefore be highlighted only to ascertain whether the moral conclusions I am drawing about fictional-x_i on a given occasion—say, in relation to meaning or motivation—are equally applicable across different media or whether they are medium-specific.

IMMORAL FICTION?

According to the Australian classification guidelines for films and computer games, "adults should be able to read, hear and see what they want," which, of course, includes fictional immorality. At the same time, the board recognizes that this freedom must be weighed against community concerns about depictions that appear to condone or incite violence, particularly sexual violence.

Anticipating such community concerns, in 2011, the Australian Classification Review Board (ACRB) refused classification of the film *The Human Centipede 2*. The ACRB effectively banned the film because, in their opinion, it contains "gratuitous, exploitative or offensive depictions of violence with a very high degree of impact." The British Board of Film Classification (BBFC) likewise refused to issue a classification because (again, in their opinion) *The Human Centipede 2* focuses on "the sexual arousal of the central character at both the idea and the spectacle of the total degradation, humiliation, mutilation, torture and murder of his naked victims" (see Shoard, 2011). Around the same time, *A Serbian Film* was also refused classification by the ACRB and BBFC (as well as others).[2] The ACRB concluded, "*A Serbian Film* could not be accommodated within the R18+ classification as the level of depictions of sexual violence, themes of incest and depictions of child sexual abuse in the film has [*sic*] an impact which is very high and not justified by context." The then South Australian attorney general, John Rau, concurred and had this to say after watching the film: "I am strongly of the opinion that *A Serbian Film* should not be released at all. . . . Some of the scenes in the DVD are so depraved that I am not prepared to even describe them in any detail."[3]

Given that the depictions noted above are distinct from what the depictions are meant to be depictions *of* (i.e., a film depicting fictitious sexual violence does not constitute actual sexual violence[4]), and given that the examples of fictional immortality are not intended to represent or otherwise refer to historical cases, they would seem to fit comfortably within my strict definition of pure fictional-x_i. Nevertheless, both the ACRB and BBFC's objections to *The Human Centipede II* and *A Serbian Film* support the view—let us call it the "immoral fiction" view—that depictions of fictional events, even when satisfying the strict definition of fictional-x_i, should be prohibited (*qua* censored) when they depict immorality *of a certain kind*, to *a certain degree*, presented in a *certain way* (e.g., sexual violence with high impact that is gratuitous). These factors make the immorality of what is depicted—even in pure fictional form—*sufficient*, at the very least, to censure the depiction (I leave the issue of prohibition and censorship until the final chapter) and imply moral disapproval.

The "immoral fiction" view is not restricted to films, of course. In the video game *Postal 2*, which was refused classification in Australia in 2005, as part of the game, one can set someone on fire while they are alive, douse the flames by urinating on them, before beating them to death with one's boot and a shovel. Similarly, the video game *Enzai: Falsely Accused* was refused classification owing to its depictions of sexual violence, including potential child sexual abuse. Nor are classification boards alone in intimating the "immoral fiction" view. When commenting on the "airport massacre" scene in the 2009 video game *Call of Duty: Modern Warfare 2*, the journalist Chick (2009) asks: Is this the most *disgusting* game of the year? In a similar tone, Roberts (2009) referred to the now-withdrawn Apple iPhone game *Baby Shaker* as *disgusting*. The game involved shaking a (virtual) noisy baby in order to stop it from crying, potentially shaking it until it died: an outcome represented by "X"s over the baby's eyes. In each case, it is not unreasonable to interpret the journalists' comments as a form of moral disapprobation.

The "immoral fiction" view likewise challenges *comic amoralism*: the claim that the content of a joke, in virtue of being a joke, and therefore occupying a particular fictional space, is neither moral nor immoral. To illustrate: When contrasting the joke, "What do you have if you have a lawyer buried up to their neck in sand?" (Answer: Not enough sound) with "What do you have if you have a [member of a minority group] buried up to their neck in sand?" (Answer: Not enough sand), proponents of the immoral fiction view would hold that, if comic amoralism were true, both versions of the joke would be morally neutral (Carroll, 2014). Instead, the latter version is likely to be interpreted by many as more morally contentious than the former. This is because the latter joke implies a certain kind of immorality (e.g., racism or potentially some other -ism like sexism) in conjunction with a certain degree of physical violence that is presented in a certain way (i.e., in a trivializing way that requires a dismissive recognition, even if not an endorsement, of the -ism depicted).[5]

AIMS

Given the potential for tension between the protection of individual freedoms and, as some would see it, the need to maintain community standards of decency, each of which features in the Australian classification system (as well as others: e.g., the United States); and given the protestations of certain journalists (as noted above) and the increased availability of the kinds of fictional content that typically create this tension—for example, depictions of extreme violence, including sexual violence (Huesmann, 2007)—it is the aim of this book to examine the morality of depictions of fictional immorality based on, among other things, their metaphysical status, content, and meaning, and the impact they have on the individual and wider society. It is hoped that a methodical examination of fictional-x_i will help inform future debate on prohibitive legislation by locating it within a cogent moral framework, supported by empirical findings and theory construction. The monograph is therefore dedicated to addressing the following question: What are the necessary and sufficient conditions for the moral condemnation of depictions of fictional immorality and our interactions with them? Or, put another way: Under what circumstances does fictional immorality count as immoral fiction?

The identification of necessary and sufficient conditions for moral condemnation will focus on (although not exclusively) the following factors:

1. The *content* of the fiction (i.e., what the depiction is a depiction of)
2. The *meaning* of the depiction within the context of the fiction taken as a whole
3. The subject's *motivation* for engaging with the fiction
4. The *medium* in which the depiction occurs (*inter alia*, literature, film, video game, and private fantasy)

In order to examine these factors, as much as possible, I will draw on genuine examples of fictional immorality (e.g., the murder of Nancy in Charles Dickens' *Oliver Twist* or sexual assault as depicted in the Japanese video game *RapeLay*) to test the merits of an argument for or against a particular factor's status as necessary and/or sufficient for a charge of immorality. I will, however, supplement these, on occasion, with fictitious examples.

It should also be noted that I intend to adopt an expressivist view of morality known as constructive ecumenical expressivism (CEE). Where immoral conditions are identified with regard to, say, the content or meaning of fictional-x_i, it is my contention that they express our negative attitude toward the object or event deemed immoral, rather than pick out some absolute moral truth. It is my further contention that this negative attitude stems from the belief that the object/event judged to be immoral realizes a property of which one disapproves; and where this attitude is shared by a community (although not necessarily for the same reason, as I shall discuss), an intersubjective or objectified moral norm is established (constructed), such that what one holds to be immoral is what one *ought* to judge immoral. Adopting such an antirealist position has implications, and potential complications, when claiming that fictional immorality is in poor taste rather than immoral, and also for censorship (again, as I shall discuss and attempt to resolve).

A TASTE OF THINGS TO COME

In this section, I briefly outline the factors identified above as candidate conditions for the immorality of fictional immorality. With regard to the first factor—content—consider the following moral judgments (M1 to M3) and the relationship on which each is based:

M1: Because what is depicted by fictional-x_i is a depiction of immorality (x_i), *all* cases of fictional-x_i are immoral.
M2: Even though immorality (*qua* x_i) is depicted by fictional-x_i, this is not reason enough for *any* instance of fictional-x_i to be immoral.
M3: Whether fictional-x_i is immoral or not depends on the type of immorality (*qua* x_i) depicted by the fiction. Therefore, even though fictional-x_i always depicts immorality, only in *some* cases is the depiction itself, in virtue of what it depicts, immoral.

M1 judges that all depictions of fictional immorality are immoral and should therefore be morally condemned, precisely because they depict immorality. Depicting immorality in the context of fiction is therefore *sufficient* for moral condemnation. M1 is likely to be considered too restrictive, however, and certainly does not approximate to the "current state of play": that is, the way our society typically views depictions of fictional immorality. M2, in contrast, states that, in the context of fiction, the relationship between the depiction and what it is a depiction of has no bearing on the moral status of the depiction itself. This fact does not negate the possibility that another reason exists to admonish the depiction; it merely denies that a sufficient reason for declaring that the depiction is immoral is that it depicts that which is immoral. While M2 allows the possibility that other factors (yet to be discussed) are available to condemn depictions of immorality, it is nevertheless likely to be viewed by some (many?) as too lenient, given that it does not place any constraint on depictions of fictional immorality based on the fact that they *depict* immorality. Again, this falls short of the "current state of play," which, typically, does not endorse an "anything goes" policy for content (i.e., in many countries virtual pedophilia *qua* child pornography is both morally condemned and illegal; see Al-Alosi, 2018, for a detailed discussion). Finally, M3 advocates *selective condemnation*, whereby the fact that a depiction is depicting immorality is sufficient to admonish the depiction in *some* (yet to be determined) cases but not others. It all depends on what is being depicted. Of the three, M3 perhaps aligns more closely with the "current state of play" regarding societal views (including a number of classification board guidelines) on the morality of fictional immorality.

Suggesting a closer match between M3 and the current state of play is not, however, the same as justifying M3's normative credentials. In other words, it is not the same as accounting for why we ought to adopt M3 when considering the moral status of depictions of immorality; nor does it provide insight into which depictions should be admonished and which should not. More work therefore needs to be done if we are to endorse M3 as our normative standard bearer.

Such work is precisely what this book intends to carry out. Moreover, as well as the *content* of fictional representations, other potential candidates need to be assessed in order to ascertain whether they, too, are necessary and/or sufficient to condemn morally depictions of fictional immorality. Regarding the second factor listed above—namely, the *meaning* of a particular depiction within the context of the fictional narrative taken as a whole (*qua* immoral worldview, or IWV)—consider the following:

> IWV: Where a fiction, taken as a whole, endorses an immoral worldview, then any depiction of immorality from within the narrative, congruent with that worldview, is immoral.

In accordance with IWV, where an immoral worldview is promoted through the fictional narrative, then depictions from within it conducive to that endorsement are immoral. The video game *Ethnic Cleansing*[6] is a strong candidate for a fiction that satisfies IWV. This is because the immoral worldview depicted within the video game, and allegedly endorsed by it, is that favored by white supremacists (see Left, 2002). Where the meaning taken as a whole is less clear, however, then the extent to which this factor is satisfied (i.e., immoral meaning) is open to interpretation. To illustrate: Should the violence in *Grand Theft Auto V* be viewed as satirical? Is it trivializing real violence (irrespective of satire), and/or perhaps even endorsing violence by inadvertently encouraging us to delight in the idea of it? Uncertainty over these questions means that, while the condition set be IWV may be acceptable, in principle, the *ease* with which IWV can be applied to films, video games, or other media used to depict fictional immorality is another matter. Again, this is something that I intend to examine further as we progress.

As for the third factor—the subject's *motivation* for engaging with fictional immorality—we have already touched on *a* potential motivation for this. With the video game *Ethnic Cleansing*, it may be that S engages with the fiction because he believes that the gameplay endorses a worldview he supports. S therefore either wishes to enact how he sees reality or, perhaps more likely, how he believes reality ought to be. Of course, it does not follow, necessarily, that S plays *Ethnic Cleansing* for either of *these* reasons. Nevertheless, our intuitions may lead us to draw just such a conclusion when confronted with certain forms of fictional immorality compared to others. To illustrate: Suppose S wishes to play Stephanie Patridge's fictitious video game *Child Sexual Assault* (Patridge, 2013) or the Japanese rape game *RapeLay*, or to watch the unlicensed *A Serbian Film*, which, you will recall, former South Australian attorney general, John Rau, described as containing "depraved" scenes of sexual violence, some of which target children. After expressing the desire to engage with these, or similar, fictions, a popular retort might be: "Why would you want to do *that*?" Such a retort would likely fail to hide the incredulity with which it was asked. An incredulity that is suggestive of a nefarious motivation: that S has an immoral reason for wanting to engage with this form of fictional immorality, and even that there is something "not quite right" with S for wanting to do this. Such a view is likely to be reinforced if S were also to declare that they *enjoy*—as in, find thrilling or are amused by—engaging with this and similar sorts of fiction. One's incredulity would perhaps be based on the following intuitions:

(a) That it is immoral to enjoy *these* depictions
(b) That there must be something wrong with someone who *enjoys* these depictions

What is implied by (b) is that S enjoys depictions of, for example, fictional rape and pedophilia because S is morally corrupt. Or, put differently, S would have to be morally corrupt to enjoy these depictions. Of course, it may be that any moral corruption from which S suffers (if this is indeed the case) is coincidental and has no bearing on his enjoyment of fictional rape or pedophilia, or motivation for engaging with fictions like *RapeLay* or *A Serbian Film*, and so on; but this is not where the intuition is leading us. Instead, the intuitive pull—that there is a causal connection between one's moral corruption and one's enjoyment—is based on the further, antecedent, intuition that such depictions are immoral, otherwise (and with reference to (a)) why would it be immoral to enjoy that which is *not* immoral?

Earlier, I stated that, with the exception of private fantasy, I intend to treat different media as part of a homogenous group unless otherwise stated. In light of this fact, before proceeding, a caveat is required. The grouping together of individuals who wish to watch, and may even enjoy watching, *A Serbian Film* with those who are willing to play and may even enjoy playing *RapeLay* (for example) does not necessitate a common motivation, despite any reference to an alleged intuition to that effect. While it is true that these examples of fictional immorality possess certain similarities with regard to content (that might turn out to be quite superficial similarities) such as depictions of rape and possibly pedophilia, they nevertheless differ in other ways that may prove to be significant when examining motivation(s) for engaging with the fiction. One obvious example is that *RapeLay*, unlike *A Serbian Film*, is a game that includes, as a feature of the game, the enactment of rape; while this may not be mandatory, it appears to be a major ludic and narrational requirement. I will return to this point below.

I also accept that enjoyment is not necessarily someone's sole reason or, in fact, *a* reason for engaging in fictional immorality, whether of the kind depicted in fictions such as *RapeLay* and *A Serbian Film* or something else. As such, there is much that still needs to be unpacked on the subject of motivation, given that I have thus far conducted only the briefest excursion into the topic. I will therefore return to the morality of one's motivation to engage with fictional immorality periodically as the monograph proceeds but focus on enjoyment in a later chapter.

In relation to the fourth factor—which is concerned with the medium used to depict the fictional immorality—I will consider whether the fact that fictional-x_i is depicted within a space intended for *playful* interactions, such as a video *game*, as opposed to, say, a film or a theater production in which, typically, one is invited to adopt or at least consider a particular moral stance while being entertained, should make a difference to our moral judgment. To be clear, it is not that one cannot adopt a moral stance while engaged in play or that one is never invited to do this (see Sicart, 2009); rather, it is that it seems to be more of a critical requirement in the case of fiction not constitutive of play (Kreider, 2008; Poole, 1982). Again, the potential differences between a film like *A Serbian Film* and the video game, *RapeLay*, are relevant here.

Finally, interwoven within all of this discussion will be an examination of the aesthetic quality of the fictional immorality under scrutiny and the extent to which aesthetic appreciation (among other things) mitigates moral concern: whether in relation to content, meaning, or one's motivation to engage with it. To illustrate: "Piss Christ" refers to artist Andres Serrano's photograph of a statuette of the crucified Christ immersed in a vat of the artist's own urine, produced in 1987 and exhibited under the title *Immersions (Piss Christ)*. For some, the image is extremely offensive, perhaps even immoral or obscene (see Young, 2000). For others, it is none of these things. Decades after its creation, debate continues over the extent to which the artwork's aesthetic properties (among other things) mitigate its alleged moral failings (see, for example, D. Casey, 2000; M. Casey et al., 2000; Holpuch, 2012). Similarly, when the Jewish Museum in New York exhibited a series of contemporary artists' representations of the Holocaust, as part of its *Mirroring Evil* exhibit, protests ensued (Biber, 2009). An eighty-one-year-old survivor of Buchenwald concentration camp, Issac Leo Kram, carried a placard on which was written, "I was there. I testify: Genocide is not art!" (Kershaw, 2002; cited in Biber, 2009, p. 227).[7] Might the aesthetic qualities of the depictions of violence in, for example, Stanley Kubrick's *A Clockwork Orange* or Quentin Tarantino's *Kill Bill* (vol. 1 & 2) mitigate claims of immoral fiction? As Symonds (2008) notes: "The body in motion in martial arts can be as beautiful as it is violent" (p. 150). The relationship between aesthetics and violence, *qua* the aestheticization of *violence* (see Appelbaum, 2017; Symonds, 2008 for recent detailed examinations) or other forms of immorality, will be taken up throughout the monograph in conjunction with factors already mentioned (such as content, meaning, and motivation).

In concluding this section, it is also evident that engaging with depictions of immorality holds a certain allure for a number of people (Schulzke, 2011; see also Konijn & Hoorn, 2005), whether in the more passive role of watching a violent thriller or slasher film or as a more active agent within a "violent" video game like *Grand Theft Auto*, in which the unfolding of events is often causally dependent on the decisions one makes within the game. In the case of video games, players can engage with different emotions and identities much more readily than when merely observing the narrative unfold in the case of, say, a film, book, or play and so invest in their own form of moral and psychological exploration (Tavinor, 2005; Young & Whitty, 2011).

When one bludgeons a stranger to death with a kitchen utensil, as it is possible to do in *Manhunt 2* (for example), or shoots a drug dealer in the face in *Heavy Rain*, one is enabled through these games—through these examples of fictional immorality—to become the architect of one's own disgust (Jansz, 2005). Within the media and academia, however, opinions differ over whether such action is morally wrong and psychologically unhealthy (Ferguson, 2008; Nauroth et al., 2014; Sjöström et al., 2013).

The empirically fuelled debate on the effects of fictional violence—whether portrayed (*inter alia*) in comic books, particularly Japanese *Manga*, or in films or video games—is presently stagnating with no consensus forthcoming on their harmful effects (see Appel, 2011; Ferguson, 2009, 2013; Kirsh & Olczak, 2002; Tan & Scruggs, 1980). In response to the lack of consensus (which is made particularly evident in recent discussion on the findings of research on "violent" video games), through a combination of philosophical analysis, and the integration of empirical research and theory construction from fields such as psychology, aesthetics, media, and games studies, it is the ultimate aim of this book to present a moral framework on which to ground any future debate on the nature and extent of prohibitive legislation relating to depictions of fictional immorality. The purpose of this book is therefore to cultivate theoretical advancement through a synthesis of philosophical analysis and existing empirical-based knowledge and theory, in order to provide a novel and comprehensive moral framework geared toward the depiction of fictional immorality and our engagement with it.

To that end, I aim to defend the view that fictional-x_i is immoral *if* it depicts immorality congruent with an immoral worldview that is being endorsed by the fiction, or *if* engaging with the fiction (*qua* one's fictive imagination) causes a certain type of harm (to be discussed), or is created for the purpose of disrespecting an actual person's inherent value, irrespective of harm caused, or *if* one uses the fiction to elicit a certain type of enjoyment (again, to be discussed). While these are not the only claims I will examine and defend—I will, for example, show that there are no nontrivial necessary conditions for fictional-x_i to be judged immoral—they are all sufficient conditions for the *immorality* of fictional immortality. Moreover, given my endorsement of CEE, I will argue not only that the conditions for immorality noted above are compatible with CEE but, importantly, that declaring that fictional-x_i is immoral is not sufficient to justify censoring it. It is not sufficient because censorship would hinder CEE's ability to construct what I will show are *robust* moral norms and, for this reason, undermine CEE as a system of morality. The only exception to this position is where a lack of censorship prevents at least one of the conditions necessary for freedom of expression to be realized: for these conditions enable CEE to work effectively.

Finally, a further important point to note and emphasize is that my interest is in the *morality* of fictional immorality. As such, while I recognize, and intend to comment on, the potential for the aesthetics of a fiction to mitigate accusations of immorality, I have nothing to say about whether a fiction's moral flaws enhance or diminish its aesthetic quality (see Gaut, 2007, for a detailed discussion on this topic).

CHAPTER OVERVIEWS

Chapter 2: The Metaphysics of Fictional Immorality. In this chapter I explore the nature of fictional existence, particularly in its "pure" form, as set out in chapter 1. I consider questions such as: In what sense can it be said that a fictional character exists, and what sort of things can be true or false *of* or *about* fictional entities? Responses to these questions ground the moral examination of fictional immorality that follows.

Chapter 3: The Content of Fictional Immorality. In this chapter I examine whether the type of immorality depicted has any bearing on its moral status and therefore whether the content of fictional immorality is sufficient, in at least some cases, to render it immoral. I argue that content alone is not sufficient for a depiction of pure fictional immorality to be immoral, even in cases of fictional depictions of pedophilia—something that is illegal in a number of countries—or where the fiction includes the use of actual racial, or otherwise discriminatory, slurs.

Chapter 4: The Meaning of Fictional Immorality. In this chapter I defend IWV (immoral worldview; see chapter 1) but acknowledge that it would be difficult to implement its conditions owing to the likelihood that different interpretations of the same fiction exist at any given time. I therefore consider an amended version of IWV based on a *reasonable interpretation* of the fiction but reject this because it is too harsh, given that it is based on an *interpretation* of the fiction's meaning. Instead, I favor a further revision that allows, minimally, that a depiction of fictional immorality is *morally insensitive* if a reasonable interpretation of the fiction cannot establish whether it is merely depicting, rather than endorsing, an immoral worldview. I also argue that moral edification and the cultivation of aesthetic appreciation have the potential to mitigate a charge of moral insensitivity.

Chapter 5: The Harm of Fictional Immorality. In this chapter I defend the claim that engaging with fictional immorality is immoral if it causes harm. In defense of this claim, I discuss what makes harm immoral, as well as what it means for someone to cause or suffer harm. I then discuss ways in which engaging with depictions of fictional immorality might satisfy these conditions, including a precautionary approach to harm. Throughout the chapter, I avoid discussing the findings of empirical research looking at the effects of "media violence" on behavior, cognition, and affect. This is reserved for chapter 6.

Chapter 6: Is There Evidence of Harm? In this chapter I present an overview of research findings looking at the alleged harmful effects of depictions of "media violence" on behavior, cognition, and affect. Owing to inconsistent findings and differing interpretations, I argue that there is a lack of consensus over the effects of engaging with depictions of fictional immorality. As a consequence, the case for harm, while sufficient, in principle, for the immorality of engaging with fictional-x_i, currently lacks consistent empirical support.

Chapter 7: Enjoying Fictional Immorality. In this chapter I focus on whether it is immoral to enjoy depictions of fictional immorality. I argue that whether it is or not depends not only on the type of enjoyment one expresses but, also, on a distinction between enjoying the depiction itself and what the depiction is a depiction of. Only in the latter case—where one uses the depiction as a means of enjoying, vicariously, that which the depiction represents (e.g., actual murder or rape or pedophilia, and so on)—should one's enjoyment of fictional-x_i be deemed immoral.

Chapter 8: Resisting Fictional Immorality. In this chapter I consider the phenomenon of imaginative resistance and whether our unwillingness to engage with certain depictions of fictional immorality should be used to inform decisions about the morality of fictional-x_i. I argue that none of the dominant explanations of imaginative resistance is sufficiently robust to proffer a normative position and, therefore, at best, the phenomenon provides insight into our own tolerances, only.

Chapter 9: Poor Taste and Fictional Immorality. In this chapter I examine what is involved in exhibiting poor *moral* taste, and therefore what is required for one's engagement with fictional-x_i to be deemed "in poor taste." I argue that accusations of poor taste express one's negative attitude toward the (perceived) treatment of something one holds to be morally pertinent, such as discrimination or murder. I also argue that, in the case of poor taste, the wrongdoing intimated by the pronouncement "*That's* in poor taste" is equivalent to a suberogatory action and is not therefore something we are morally *obliged* not to do.

Chapter 10: Historical Fiction and Fictional Immorality. In this chapter I examine the extent to which the morality of fictional immorality is affected by the fact that it features (for example) actual historical figures engaged in fictional immorality (e.g., Italian composer Salieri murdering Mozart in the film *Amadeus*). While published discussion on historical fiction often focuses on the extent to which what is depicted strives (successfully or not) to be historically accurate and/or experientially authentic, such concern is not relevant to cases of fictional immorality, as the *fictional* aspect of the depiction is deliberate. Because of this, concern over the morality of historical fiction (so described), for the most part, matches previous discussion in relation to pure fictional immorality. An exception to this claim concerns damage to the actual person/historical figure's reputation. Such concern is, however, dependent on the extent to which a particular narrative involving fictional immorality is being endorsed rather than presented as make-believe or where the possibility of endorsement is ambiguous—again, as discussed in previous chapters.

Chapter 11: A New Kind of Fiction. In this chapter I consider the morality of a relatively new and specific type of fiction that, again, depicts actual persons: namely deepfake pornography. I argue that deepfake pornography is typically a form of nonconsensual pornography and image-based sexual abuse. It is therefore immoral for these reasons. I also argue that the manufacture of deepfake pornography is immoral because it is disrespectful, minimally, to the target of the deepfake manipulation.

Chapter 12: Fantasy and Fictional Immorality. In this chapter I assess the morality of *private* fantasy as a distinct category of fictive imagination and consider whether the fact that one's private fantasies are only for one's own consumption protects them from certain, perhaps *all*, moral admonishment regardless of content, meaning and motivation. In other words, I examine whether the privacy element, unique to this particular means of depicting fictional immorality, makes moral concerns about private fantasy a nonstarter. It is my contention that it does not, for reasons relating to disrespect. I do, however, argue that certain types of fantasy (e.g., idle fantasy) could amount to a lesser (suberogatory) wrong, even when they contain similar depiction to those found in pornographic deepfakes.

Chapter 13: "It's Not Immoral, but It Is in Poor Taste." In this chapter I defend a metaethical account of morality known as *CEE*. According to CEE, declaring that "*x* is immoral" is equivalent to expressing a negative attitude toward a property (P) one believes is realized by *x*. The means by which one arrives at judgments about the morality of fictional immorality, *mutatis mutandis*, appears to be the same as that ascribed to cases of poor taste (see chapter 9). Given this, in this chapter, I develop the argument first presented in chapter 9 that poor taste concerns one's attitude toward the (perceived) *treatment* of a morally pertinent matter, thereby making poor taste parasitic on immorality. The implications for fictional-x_i of the relationship between poor taste and immorality are then explored.

Chapter 14: Immoral Fiction and Censorship. In this final chapter, I summarize the preceding chapters' key points, focusing on the sufficient conditions for the immorality of fictional-x_i and one's engagement with it. The summation is intended to provide a framework for understanding the morality of depictions of fictional immorality. After that, I discuss the relationship between morally problematic fiction (whether immoral or in poor taste) and censorship, particularly from the perspective of CEE. I conclude that the immorality or poor taste of fictional-x_i is not sufficient for it to be censored.

NOTES

1. In this respect, my use of the term "fiction" is much narrower than that of Walton (1990), for example.
2. The BBFC did approve an edited version for cinema and DVD release. An edited version was submitted to the ACRB, who still refused approval.

3. See also https://www.independent.co.uk/arts-entertainment/films/features/a-serbian-film-is-this-the-nastiest-film-ever-made-2137781.html (accessed July 25, 2019).

4. In relation to this point, it is worth noting that, in the UK, a charge of gross indecency was brought against Howard Breton for a simulated homosexual rape scene in his 1980 play, *Romans in Britain*. The point of contention was whether simulating an act of gross indecency (a legal term) is itself a grossly indecent thing to do. It is also worth noting that the prosecution was eventually withdrawn. https://www.theguardian.com/world/2015/mar/19/romans-in-britain-play-mary-whitehouse-trial-archive-1982 (accessed January 23, 2020).

5. What humor there is in either version of the joke may turn on a violation of expectation, owing to the way the joke is set up. Nevertheless, such a violation of expectation as an expression of violent racism is what is likely to cause unease. That said, I am not suggesting that racism in the absence of violence (but not violation per se) is morally acceptable.

6. Developed by Resistance Records 2002.

7. See also https://www.nytimes.com/2002/03/18/nyregion/exhibition-with-nazi-imagery-begins-run-at-jewish-museum.html (accessed December 8, 2019). It is also worth noting that I recognize that the examples of Piss Christ and the Holocaust do not fit my strict definition of pure fictional immorality. They are, however, applicable to discussion to come on historical fiction.

Chapter 2

The Metaphysics of Fictional Immorality

> *Of all bad deeds that, under cover of the darkness, have been committed within wide London's bounds since night hung over it, that was the worst. Of all the horrors that rose with an ill scent upon the morning air, that was the foulest and most cruel.*
>
> —Charles Dickens, *Oliver Twist*, 1838

INTRODUCTION

The extract from *Oliver Twist* alludes to the murder of Nancy by Bill Sikes. In chapter 1, I referred to this as an example of pure fictional immorality, whereby fictional-x_i depicts an immoral action (x_i) that is not directed toward, nor does it involve in any significant way, the depiction of an actual object, animal, person, or event or otherwise allude to any of these. Bill Sikes is a fictional character, as is his victim, Nancy. Moreover, the murderous act depicted by Dickens is a fictitious act. Prima facie, the requirement for pure fictional-x_i appears to have been met, as neither the characters nor the immoral event depicted satisfy the criterion of depicting that which is "actual." At the time of their creation, Bill and Nancy were not intended to represent anyone who actually existed—in a *robust* sense (to be explained)—or who had once existed (in the same robust sense). *Mutatis Mutandis*, the same can be said of Dickens' depiction of the murderous event.[1]

A level of irrealism is associated with the kinds of fictional characters described that accord with our pretheoretical intuitions or folk view of fictional entities: that fictional characters do not exist (Zvolenszky, 2013). For proponents of the nonpredication theory of literature (Stern, 1965), this means that the proposition “Nancy was murdered” and the subsequent claim that Bill Sikes murdered her are neither true nor false, because a presupposition of the depiction of the murderous event is not taken to be true. In other words, the event described by Dickens can be neither true nor false because that which would otherwise be taken for granted if the event described had been actual—the truth that the characters depicted exist or existed, and the event described occurred—is understood not to be true (Kajtár, 2017). Yet there is clearly a sense in which the proposition “Fagin murdered Nancy” is false and the proposition “Bill Sikes murdered Nancy” is true (Friend, 2007; Grittenden, 1966; Predelli, 1997). (Imagine, for example, answering the quiz question: Who murdered Nancy in the novel *Oliver Twist*?) Moreover, it is far from ridiculous to assert that Sherlock Holmes is the most famous detective in the world—a claim that is likely to receive empirical support—even when admitting moments later that Sherlock Holmes is not an actual detective.

Given this, what is the metaphysical status of pure fictional-x_i, and the fictional characters directly involved in the immoral fiction, that makes it possible, in one sense, for the propositions “Bill Sikes is a murderer” and “Sherlock Holmes lives at 221b Baker Street” to be neither true nor false and yet, in another, quite clearly true? Not only is such a question germane to examining the moral status of pure fictional immorality; I would go so far as to say that answering it is a prerequisite. For it is important to determine what sorts of moral judgments can be made, and with what degree of confidence, toward a target that apparently lacks truth-aptness in some contexts but not all and whose nature is therefore potentially unclear. Put differently: What sense is to be made of the claim that fictional-x_i is immoral where fictional-x_i is neither true nor false given that it depicts that which is not actual? In posing this question, I am not suggesting that there can never be a coherent claim made about the immorality of fictional immorality, only that the metaphysical nature of fictional-x_i, and therefore the depiction of murder in Dickens’ *Oliver Twist*—as a paradigm example of fictional immorality—needs to be clarified before a compelling argument can be delivered either way.

A Nominal Commitment

It is worth noting from the outset that in describing Bill and Nancy as fictional characters, I am committing myself to a certain ontological position, albeit a fairly conventional one: that there are such things as fictional characters. Consequently, at a nominal level, I consider it appropriate, ontologically speaking, to distinguish between fictional and nonfictional entities, and therefore (inter alia) between fictional and nonfictional characters and persons. Thus far, there is nothing remarkable about what I have said or committed myself to. Nevertheless, from a metaphysical perspective, the following ontological assertion needs greater clarification: "if there are such things as fictional entities, characters, persons, then fictional entities, characters, persons *exist*" (Meinong, 1904/1960; van Inwagen, 1983). The remainder of the chapter is dedicated to just such a task, and therefore to addressing the metaphysical question: What sort of existence is fictional existence? To reiterate, the ensuing discussion should not be viewed as a mere academic exercise but as prerequisite to, and therefore the necessary basis for, the moral examination of fictional immorality to come.

In what is to follow, I will focus on fictional characters as fictional persons (*qua* human beings) and will use the terms interchangeably. I do so with the understanding that while all fictional persons are fictional characters, not all fictional characters are fictional persons (e.g., Shere Khan from Rudyard Kipling's *Jungle Book*, or dementors as featured in *The Prisoner of Azkaban* and other stories by J.K. Rowling). I also take it as a given that fictional characters and persons can be subsumed under the more general term "fictional entity."

THE CONDITIONS FOR EXISTENCE

The proposition "All triangles have three sides" is necessarily true. The proposition "ϕ thinks" is not (Ayer, 1953; Abraham, 1974); instead, its truth is dependent on performance (Austin, 1962; Hintikka, 1962). In other words, the proposition "ϕ thinks" is true *if* (and therefore *because*) ϕ thinks or asserts it (perhaps while entertaining the possibility that it is false). Moreover, if ϕ thinks, then ϕ exists. ϕ's act of thinking is sufficient for ϕ's existence, although the exact nature of this existence has yet to be determined. It is also necessary if ϕ's nature is exclusively that of a *res cogitans* (a thinking thing). Under such circumstances, the following would hold: ϕ exists *if* and *only if* ϕ thinks. But if ϕ is not a *res cogitans*, then the previous proposition does not hold. From this, we can conclude that the conditions for existence differ depending on whether ϕ is a *res cogitans* or not.

The nomological truth of a *res cogitans* is a contentious issue; nevertheless, it remains a metaphysical possibility. It is included, here, to illustrate the lack of universal conditions for existence. Owing to a lack of universality, I am—without fear of contradiction—able to commit, nominally, to the existence of fictional persons while acknowledging the legitimacy of the claim that the fictional characters Sherlock Holmes, Victor Frankenstein, and Homer Simpson do not exist, have not existed, and will not exist (Sanson, 2016). This is because a *nominal* commitment requires nothing more than the recognition that fictional characters satisfy certain existential conditions but not others, and therefore they qualify for one type of metaphysically possible existence but not another. As things stand, then, the term "existence," at least as bandied around thus far, is vulnerable to the charge of equivocation.

As we progress, I intend to increase my existential commitment to fictional persons (as well as fictional entities more generally), beyond the current nominal level, by delineating the nature of fictional existence in terms of necessary and sufficient conditions. In the meantime, I begin my defense of a nominal-level commitment by reiterating the fact that a fictional person differs from a *res cogitans* at least insofar as the former is not a thinking thing. Therefore, the truth of the proposition "ϕ thinks" is neither necessary nor sufficient for the existence of a fictional person. Moreover, fictional persons differ from actual persons (*qua* human beings) at least insofar as the former is noncorporeal. Corporeality is therefore neither necessary nor sufficient for the existence of fictional persons. Given these differences, the following universal claim about existence (U_e) does not hold:

- U_e: ϕ exists *iff* ϕ satisfies the conditions for existence[2]

U_e does not hold because it implies a set of universal conditions for existence that ϕ as a *res cogitans* is intended to repudiate, owing to the fact that the set of existential conditions for a *res cogitans* differs from those required for other entities that incontrovertibly exist (such as, rocks, trees, human beings). In the place of U_e, I insert the relativistic proposition, R_e (relative existence), although I do so tentatively because I concede that R_e is little more than a truism. Its relevance to the discussion will become apparent as we continue, however.

- R_e: ϕ exists *iff* ϕ satisfies the conditions of ϕ's existence

In and of itself, R_e tells us nothing that we could not already deduce from the existence of ϕ: for the fact that ϕ must satisfy the conditions of ϕ's existence is part of what it is to understand the truth of the proposition "ϕ exists." How else could the proposition "ϕ exists" be true unless the conditions of ϕ's existence had been met? What is missing from R_e is an indication of what these conditions are, *relative to* ϕ. Leaving this issue aside for the moment, what we can say, thus far, is that the necessary and sufficient conditions for the existence of ϕ, *qua* a *res cogitans*, are that it thinks and that this is not an existential requirement for something that is not a *res cogitans*, such as a table or rock, or tree, amoeba, or fictional character. All *non-res cogitans* exist without thinking: meaning that the act of thinking is neither necessary nor sufficient to bring any of them into existence. The fact that they do exist (at least at the nominal level of existence I am operating at here) indicates that all of the conditions for *their respective existences* have been met, and none of these includes the act of thinking.[3] Therefore, for the existence of each metaphysically possible entity to occur, there cannot be a shared set of conditions for existence. Thus, if one were to inquire "Does Cartman from *South Park* exist?," while the following response may be valid—namely, "Cartman from *South Park* exists *if* and *only if* Cartman from *South Park* satisfies the conditions of Cartman from *South Park's* existence"—it is, needless to say, informationally vacuous.[4] The next step, then, is to identify the necessary and sufficient conditions for the existence of fictional characters.

FICTIONAL EXISTENCE

Abstract Entities

Thus far, we can say that thinking, corporeality, and tangibility are neither necessary nor sufficient for the existence of fictional characters.[5] Fictional characters exist, instead, in a manner Mole (2009) describes as *ontologically undemanding*, such that if one were to provide a metaphysical or scientific account of what exists in *our* world, it would not include Bill Sikes or Nancy, nor any other fictional entity. Put differently, if one were to list all of the known victims of homicide named Nancy throughout the history of the world, Dickens' creation would not (should not) be among them (Yagisawa, 2001). To exist in an ontologically undemanding way is to exist in a manner that aligns itself with our pretheoretical intuition that Nancy and Bill, along with every other fictional character, do not exist in a robust sense: that is, exist, at the very least, as *tangible* entities or *concreta*. Rather, they exist (inter alia) as nonthinking, noncorporeal, nontangible entities, and *only* in their corresponding fictional world.

Abstract entities or objects (Quine, 1960)—for example, numbers, ideologies (such as "justice for all" or "animal rights"), institutions (such as marriage)—are (inter alia) nonthinking, noncorporeal, nontangible entities, identified, here, by *way of negation* (Lewis, 1986): that is, in terms of what they are not or what they lack. We may agree with the ideology underpinning animal rights, for example, or same-sex marriage, or consider one or both to be misguided; but in neither case do we typically think of them as fictions. Therefore, if fictional entities are abstract objects, they must be of a kind that can be distinguished from nonfictional abstract objects (Livingston & Sauchelli, 2011). One distinguishing feature is that certain abstract objects arguably exist independently of our existence. Uncontroversial examples are numbers or geometric shapes (*Platonic abstracta*). Far more contentious is the claim that moral facts pertaining to what is good or bad, right and wrong, exist independently (Moore, 1903). In the case of numbers, given their generally accepted independence, we are not typically believed to have *created* numbers but, rather, discovered them and their mathematical relations. In contrast, whether established moral facts or truths (such as "murder is wrong") were created (*qua* socially constructed) or discovered is much more contentious (Harman & Jarvis Thomson, 1996).

In the case of fiction, it is common practice to say of an author that they *created* the fictional character, rather than discovered it. It would be somewhat unorthodox to say of Sir Arthur Conan Doyle that he stumbled across Sherlock Holmes. Certainly, such a claim would require clarification, with the likely expectation that what is explained is how the author *created* rather than discovered the character in the allegedly fortuitous way intimated. Yet it would not be strange for the reader to say that they had stumbled across the story and therefore the character, Holmes, by chance. What this tells us is that, once created, the fictional character has some existential independence: Sherlock Holmes is "out there," waiting to be discovered by the naïve reader who happens upon, say, a copy of *A Study in Scarlet* or *The sign of Four*. But, equally, one should not exaggerate the independence of fictional entities by concluding that they exist in a manner consistent with numbers, mathematical relationships, or geometric shapes: for they do not. It would therefore be incorrect to claim that Sherlock Holmes would continue to exist if no human being or an equivalent literary community existed that was capable of grasping the fictional status of Conan Doyle's creation.

The existential position described is reminiscent of *creationism* about fictional characters, also known as *abstract artefact theory* (see Thomasson, 1999, for a detailed discussion). According to creationism, fictional characters are the product of the human mind (Vecsey, 2015) or, more specifically, contingently dependent for their creation and continued existence on the intentions of the author *and* a literary community.[6] Thus, for proponents of creationism, once created, the fictional character would cease to exist at the point where the literary community likewise ceased to exist. Indeed, one is left to speculate about the number of fictional characters lost to posterity following the demise of a particular oral tradition (Leclerc, 2016). Of course, many fictional characters are contained within books or other media and are therefore not dependent for their existence on a community that engages exclusively in storytelling as an oral tradition (imagine, for example, if the Anglo-Saxon poem *Beowulf* had not been committed to manuscript around the tenth and eleventh centuries CE). Nevertheless, if a literary community capable of discerning the fiction no longer exists—to read or otherwise engage with it—then the fictional character will cease to be. In short, creationism accords with our pretheoretical intuition and folk understanding of fiction: that fictional characters are created by the author but, importantly, exist as fiction through the appropriate actions of the literary community.

The interdependence of author and audience will be discussed further, below. In the meantime, let us consider the nature of fictional existence in more detail.

Speech Act Theory

Suppose after reading *A Study in Scarlet*, S declares: "I wish I were a great detective like Sherlock Holmes." As plausible as the scenario is, what are we to make of it, especially given that S is aware of Holmes' fictional status? *Speech Act Theory* (Austin, 1962; Searle, 1979) is concerned with the role of language in communication. When S utters the sentence, "I wish I were a great detective like Sherlock Holmes," it may be that she is conveying regret (i.e., alas, I am not a great detective like Sherlock Holmes, but I wish I were) or aspiration (i.e., I aspire to be a great detective, just like Sherlock Holmes). Importantly, though, for the speech act to be successful—that is, for meaningful communication to occur—what S refers to as an example of a great detective—namely, *Sherlock Holmes*—must exist. If Sherlock Holmes does not exist, then the speech act's role, in meaningfully communicating regret or aspiration, fails. It would be equivalent to saying: "I wish I were a great detective like %#!"[7] For proponents of creationism, the fact that the speech act is successful at communicating the subject's regret or aspiration means that Sherlock Holmes *exists* (Searle, 1979).

But surely it is the case that, in being fictional, Holmes does not literally possess the investigative acumen S aspires to have or regrets not having. That is, Holmes does not instantiate the property of "clarity of thought," for example, or "heightened reasoning." In fact, Holmes does not instantiate any properties at all. For Leclerc (2016), this is because only *concreta* are capable of instantiating properties. To hold that such instantiation occurs in the case of fictional characters is to make a category mistake (Zvolenszky, 2013). But all of this can be true without conceding that fictional characters do not exist and therefore cannot be referenced in the context of Speech Act Theory. To understand why, it is important to note that, fortunately, it is not a requirement of Speech Act Theory that the meaningfulness of the declaration "I wish I were a great detective like Sherlock Holmes" is dependent on the existence of an entity capable of property instantiation. Instead, S (as the subject of the utterance) can simply aspire to possess properties the author of the fiction ascribes to Holmes as part of his *depiction* of the fictional detective (van Inwagen, 1977). In other words, properties Conan Doyle *describes* Holmes as possessing (either directly or indirectly in virtue of the many descriptions of the reasoning and action Holmes undertakes within the narrative), rather than properties Holmes is capable of instantiating. Thus, "I wish I were a great detective like Sherlock Holmes" translates to "I wish I possessed the same detective skills and investigative acumen as various descriptive narratives attribute to the fictional character Sherlock Holmes."

Suppose, however, S had said: "I wish I were a great detective like Frank Serpico." Given what is required to make this alternative speech act successful, and given that Frank Serpico is not a fictional character, does this mean that, in declaring that Sherlock Holmes *exists*, for proponents of creationism, Sherlock Holmes and Frank Serpico are existentially equivalent? I do not believe so. The fact that the respective sentences referring to Holmes and Serpico are meaningful sentences simply means that the existential status of Sherlock Holmes and Frank Serpico is such that each is capable of satisfying the requirements of Speech Act Theory; but this does not necessitate that each satisfies the same conditions for existence. The threshold of existence required to accord with Speech Act Theory, and therefore make a sentence meaningful, does not necessitate *uniformity* across the condition for existence.

In the case of Frank Serpico, let us allow that, as a token *concreta*, he instantiates the property of integrity; but, as S has never met Frank Serpico, his knowledge of Serpico's integrity is indirect, based (say) on published accounts of the detective's professional life. In each case, S's utterance signifying his desire to be a detective of the ilk of Sherlock Holmes or Frank Serpico, or possibly his regret at not being equivalent to either, is based on *descriptions* of characteristics attributed to each individual, irrespective of whether the respective individual is capable of instantiating these properties or has merely been ascribed them. This is sufficient to satisfy Speech Act Theory whether one is referring to an actual individual or a fictional one. Saying this, however, does not preclude the possibility that someone could actually know (be an acquaintance of) Frank Serpico and wish to be a great detective as a consequence of perceiving directly the aspirational attributes he instantiates. Such a possibility does not undermine the point made here: that the conditions for each respective existence are sufficient to satisfy the requirement of Speech Act Theory without each "existence" being equivalent.

Invention without the Intention to Deceive

Fictional entities *are* abstract objects. Abstractness is therefore necessary for the existence of fictional characters. Yet more is needed. After all, when entertaining the truth or falsehood of the proposition "I am not a number," one has to engage with abstract entities (numbers and possibly the concept "I"); but, in doing so, one is not engaging with *fiction*, as intuitively understood. To understand what is required for something to be fictional, in addition to reference to abstract entities, consider the following utterance:

- "Ann Other is a liar and a cheat."

In scenario 1, S declares this because she believes the testimony of a witness who testifies to that effect, even though the witness' testimony is false (although S does not know this). In scenario 2, S knows that what she is about to say is false but says it anyway. In scenario 3, S's utterance is part of a story she often recites which begins: "Once upon a time, there lived a woman called Ann Other . . ." In the first two scenarios, what S has to say about Ann—that she is a liar and a cheat—is untrue, but only in the second scenario is S lying. In the latter example, S intends to mislead; in the former, she does not. Moreover, in scenario 2, S has invented a state of affairs—namely, that Ann is a liar and a cheat—with the intention of misleading those she directs her utterance toward. In the third scenario, however, although what S has to say is invented, unlike scenario 2, she does not intend to mislead her audience (*qua* those she directs her utterance toward).

It would seem, then, that a necessary condition for the existence of fictional entities is *invention without the intention to deceive* (Kajtár, 2017). Fictionality therefore involves describing, without seeking to deceive, a series of events that intentionally violates what Davies (2001) calls the *fidelity constraint*: meaning that the author intentionally *fails* to describe "only events she believes to have occurred, narrated as occurring in the order in which she believes them to have occurred" (p. 264).[8] In short, she invents them. But, importantly, for the fiction to be understood and therefore embraced as fiction, the audience must recognize the author's intent. The utterance "Ann Other is a liar and a cheat" can therefore be said to describe a fictional person with fictional character traits "if and only if the author intends the audience to imagine or make-believe its content and the audience recognizes the intention of the author" (Kajtár, 2017, p. 2173; see also Weisberg, 2016). Uttering "Ann Other is a liar and a cheat" under the various conditions described also illustrates that syntax and semantics are incapable of differentiating fiction from nonfiction (Friend, 2012).

So far, then, the criteria for the existence of fictional characters stipulate that they are:

A. Abstract objects—events, characters, persons—invented by the author that violate the fidelity constraint without the intention to deceive.[9]
B. Recognized by the audience as such (even if the audience is the author: that is, if one creates a fiction "in one's head" simply for one's own amusement).

Pure Fiction

Points A and B stipulate the necessary and jointly sufficient conditions for a nonpure form of fictional existence. For the existence of *pure* fiction, a further condition must be satisfied. Consider the character Napoleon as portrayed in Tolstoy's *War and Peace*. In accordance with points A and B, *this* Napoleon exists as a fictional character in the corresponding fictional world created by Tolstoy, and only there. Thus, even if Tolstoy had published his work during (the actual) Napoleon Bonaparte's lifetime, the two Napoleons could never have met or in any way conversed. Nevertheless, the Napoleon of Tolstoy's *War and Peace* is not a further paradigm example of the pure fictional character I have in mind in the context of fictional-x_i, as Tolstoy's Napoleon was intended by Tolstoy to represent the actual historical figure. *Mutatis mutandis*, the same can be said of the fictional Napoleon that features in Dumas' *The Count of Monte Christ*, and even *Bill and Ted's Excellent Adventure* (directed by Stephen Herek).

These Napoleons per se are not examples of fictional immorality, but even if they were depicted doing something immoral and/or having an "immoral character," they would nevertheless violate the requirement of pure fictional-x_i: namely, that the fictional element is not directed toward, nor does it involve in any significant way, the depiction of an actual object, animal, person, or event, or otherwise allude to any of these.

Returning to the paradigm example of fictional-x_i found in Dickens' *Oliver Twist*, Bill and Nancy were brought into existence through an act of creative writing. Given my commitment to an undemanding ontology, however, it is my contention that what was created—what the author brought into existence—exists in a fictional world *only*, and nowhere else. Endorsing the notion of an author (or more general creator) bringing fictional characters into existence should not, therefore, be vulnerable to a charge of ontological blurring or slippage, or some other form of equivocation: for it is not my claim that, through the act of creation, Bill and Nancy become, for want of a better word, *real* or existentially equivalent to actual entities in terms of the necessary and sufficient conditions for their existence. They do not exist in a robust sense; instead, fictional characters (and other fictional entities) exist only as abstract objects: meaning that fictional entities are *not* cut from the same metaphysical cloth as actual entities.

In accordance with these claims, I seek to adopt a fictional antirealist position, whereby Tolstoy's Napoleon, Dickens' Bill and Nancy, and Rowling's Hogwarts or dementors fail to exist in the way actual persons, places, or things exist. They exist only as abstract entities, in violation of the fidelity constraint and without the intention to deceive. So understood, it is true, and recognized as true by author and audience alike, that Bill Sikes (the fictional character created by Charles Dickens) is a murderer (within the fictional realm of *Oliver Twist*) and yet neither true nor false that Bill Sikes (the fictional character created by Charles Dickens) is a murderer (in the actual world) (Button, 2012). Such a declaration is not contradictory; in the same way that it would not have been contradictory or otherwise paradoxical (in a Moorean sense) if Dickens had said of Bill Sikes: “I created him, but I do not believe that he exists” (Leclerc, 2016). The conditions for fictional existence are therefore sufficient for us to make truth claims about the properties ascribed to the fictional character, and therefore to have knowledge of the character’s attributes, so described (Schiffer, 2003)—for example, Bill Sikes is a murderer; Sherlock Holmes is a detective—while also, from an external perspective, possessing knowledge of pretense-free truths (Vecsey, 2015), such as (I suspect) “more people have heard of Sherlock Holmes than Frank Serpico.”

CAN FICTION BE ACCIDENTLY TRUE?

Imagine I had written a song about the murder of Kitty Jane MacManister, who died as a result of being thrown off the Tallahatchie Bridge. From the outset, I make it clear to all concerned that the lyrics describe a fictitious death; they are the product of my imagination. I am therefore not trying to pass off the song as a poetic description of an actual immoral event. Prima facie, this would seem to be an incontrovertible example of pure fictional immorality (or fictional-x_i).

Even the fact that the Tallahatchie Bridge used to exist need not detract from the song’s status as pure fiction. After all, it is only a requirement of pure fictional immorality that it is not directed toward, nor does it involve in *any significant way*, the depiction of an actual object, animal, person, or event, or otherwise allude to any of these things. Suppose, however, I later discover, to my surprise, that many years before the song’s creation, someone named Kitty Jane MacManister had, in fact, been murdered in exactly the way I describe in the song. What impact, if any, should this revelation have on the status of fictional-x_i as *pure* fictional immorality?

According to Currie (1990), for ϕ to be a work of fiction, it must (i) be the product of a *fictive intent*—that is, be presented as a fiction, and therefore as untrue, a pretense, rather than as a lie masquerading as the truth–and (ii) if later revealed to be true, be *accidently* true: that is, be a depiction of an *actual event* in virtue of coincidence. My example of a song depicting the fictitious murder of Kitty Jane MacManister satisfies Currie’s first clause, as well as points A and B, discussed earlier. The later revelation that an actual murder, matching the song’s description in every detail, had occurred years earlier suggests that the depiction is both a fiction and true, even if accidently so—at least, if by “true” we mean that the terms used to describe the fiction also describe what actually occurred. Importantly, though, if this is a possibility in *one* (arbitrary) case of pure fiction—that is, the one we happen to be discussing—then, where the possibility exists that a fictional event is accidently true in virtue of being a depiction that coincides with an actual event in this world, what is to stop it from being accidently true in virtue of being descriptively identical to an event that takes place in some other metaphysically possible world?

Let us allow, for example, that on *a* twin Earth, at some point in time, a resident of that planet, named Kitty Jane MacManister, was (or is going to be) murdered by being hurled off twin Earth's corresponding Tallahatchie Bridge. If we allow fictional-x_i to be accidently true in virtue of this coincidence, then, given *all* metaphysically possible worlds, what fictional-x_i depicts is necessarily going to coincide, descriptively, with some actual event in some possible world at some point in time. More formally:

1. Fictional-x_i is accidently true *if* and *only if* what fictional-x_i depicts coincides with an actual event on some possible world.
2. Given all possible worlds, fictional-x_i will necessarily coincide with an actual event on some possible world at some point in time.
3. Given 1 and 2, it is not possible for fictional-x_i not to be accidently true.

The accidental truth of fictional-x_i is problematic for my goal of examining a pure class of fictional immorality because it negates the possibility of a depiction of this kind being free of any significant association with actual events, even when the fiction was not intended to depict actual events. To avoid the problem posed by the accidental truth of a fictional depiction, I need to demonstrate that Currie's clause (ii) need not be true and, therefore, that fictional intent (which implies a lack of intent to depict the "actual") negates the possibility of fictional content being accidently true, even when descriptively identical to an actual event. This will allow pure fictional-x_i in the absence of the accidental truth of what the fiction depicts, thereby allowing a case like the fictional murder of Kitty Jane MacManister to remain an example of pure fictional immorality. In short, in order to salvage the principle of "purity" underlying fictional-x_i, I need to show (in contrast to Currie) that a *pure* fiction cannot be accidently true (I will, however, return to this issue in chapter 10 when discussing the possibility of the accidental truth of historical fiction).

The remainder of this chapter is dedicated to defending this claim. Before doing so directly, however, let us begin by considering the argument in favor of the possibility that fictional-x_i could be both a work of fiction *and* true (*qua* depicts something that is identical to what has actually occurred, is occurring, or will occur).

Currie (1990) endorses a descriptivist theory of fiction, whereby what gives a fictional term meaning—for example, the name, Kitty Jane

MacManister—given its intended reference to an abstract entity, is the collection of descriptions that the descriptive term “Kitty Jane MacManister” is able to satisfy. The name Kitty Jane MacManister is therefore an *abbreviated descriptor* for all the *descriptions* that the abstract entity Kitty Jane MacManister satisfies, such as murder victim and, let us allow, brunette, resident of Green Wood Mississippi, person who had a newt put down her back at the Carson County Picture Show, and so on (all of which, again, let us allow, is mentioned in the song lyric). Now, if, according to descriptivism, a series of descriptions is necessary and sufficient for a fictional object to have meaning—that is, to answer the question: “What does the proper name Kitty Jane MacManister mean/refer to? (answer: *this* list of descriptive terms)—then where coincidence allows that the real-life Kitty Jane MacManister also satisfies these descriptive terms, the abbreviated descriptor is at the same time accidently true insofar as it applies equally well to an actual person.[10] In order for my definition of pure fictional immorality to avoid the possibility of being accidently true, even if I were to endorse a descriptivist account of fiction, I must show how such a coincidence between fact and fiction does not have to lead to the unpalatable (at least, given my aims) outcome: namely, a pure fiction that coincidently *depicts* actual events.

As a first step toward this goal, consider a further example. Suppose, as part of a different fiction, I wrote: “The king is dead, long live the king!” Does the fact that, throughout history, numerous actual kings have died and been replaced by other actual kings undermine the fictitious status of the present content and, in particular, the king referred to within the narrative? I do not believe so. Just because the abbreviated descriptor “king” can be applied to (and be satisfied by) any number of actual kings, alive or dead, this possibility does not negate the fact that the author of the fiction has a particular king in mind, to the exclusion of all others, when employing the term within the fiction (Lamarque, 2010; Searle, 1996).

Imagine, then, that I penned a screenplay about a white middle-aged sexual predator named John Smith, who commits a series of heinous crimes without ever being caught. In the same way that the descriptor or common noun “king” can be instantiated by any number of actual kings throughout history without this fact detracting from the fact that the author of the fiction had a particular fictitious king in mind that excluded all others, so the proper nouns John Smith can be instantiated by any number of white, middle-class “John Smiths,” alive or dead without this fact detracting from the fact that the author had a fictitious John Smith in mind within the narrative to the exclusion of all others. Moreover, even if a sexual predator is uncovered with exactly the same name, physique, ethnicity, socioeconomic status, and, let us allow, age and modus operandi as the character in the play, such a set of facts makes certain events within my fiction, at best, descriptively identical to these events without making what the *fiction* was *intended to depict* true.

In accordance with descriptivism, the ascribed lust and depravity of the fictional serial sexual predator—his described actions, thoughts, feelings—are captured by the abbreviated descriptor John Smith. How might we explain the relationship, whereby what is true *of* a pure fictional character is true *of* an actual person even though what is described as true, and therefore true as a descriptor, is not a truth *about* that or any actual person?

Suppose one were to say of fictional John Smith that "*He* did *these* heinous things." One could also say the same of actual John Smith: that "*He* did *these* heinous things." What can be truthfully said *of* fictional John Smith can also be truthfully said *of* actual John Smith. The description applies equally well to both. What the third-person and demonstrative pronouns are referring to in each case, however, is different because demonstratives (and indexicals) are context-sensitive.[11] So, while what is descriptively true of fictional John Smith is descriptively true of actual John Smith—insofar as, in each case, *he* did *these* heinous things—what the sentence is about (is referring to beyond the description) differs in each case.[12] Reference to "he" in the case of fictional John Smith, while referring to "John Smith" as described in the fiction, is also referring to, insofar as it is *about*, an abstract entity that is *bereft of any other referent* (i.e., "John Smith" in the fiction is not intended to refer to an actual person). The fictional entity is also described as doing heinous things. What this amounts to is a fictional entity fictionally doing fictionally heinous things (i.e., heinous things that are not based on historical events). In the case of actual John Smith, of course, "he" picks out (is about) "John Smith," a corporeal and biological entity who happens to engage in certain actual heinous crimes. Consequently, it is not true that what the fictional description is *about* (i.e., what the abbreviated descriptor is referring to outside of the collection of descriptive terms) applies equally well to fictional and actual John Smith. The intentional lack of referent to an actual entity (i.e., the fact that the abbreviated descriptor "John Smith" in the fiction is not itself intended to refer to an actual John Smith) is an important factor in determining whether a depiction can be accidently true or not, as I intend to show in further detail when discussing historical fiction in chapter 10.

CONCLUSION

This chapter is dedicated to addressing the metaphysical question: What sort of existence is fictional existence? I have identified what I take to be the necessary and sufficient conditions for the existence of a fictional entity. ϕ is a fictional entity *if* and *only if* it is:

- An abstract object
- Intentionally created to violate the fidelity constraint without the intention to deceive
- Recognized as such by its audience

The conditions for fictional existence are themselves sufficient for us to make truth claims about properties ascribed to the fictional entity (e.g., Sherlock Holmes is right-handed) but, also, for us to have knowledge of pretence-free truths, such as (I have speculated) more people have heard of Sherlock Holmes than Frank Serpico.

Finally, under the guise of *pure* fictional immorality, the fiction must not be directed toward, or involve, in any significant way, the depiction of an actual object, animal, person, or event, or otherwise allude to these things. I discussed how the possibility of a fiction being accidently true is problematic for the notion of pure fiction and therefore pure fictional immorality, as it appears to negate their existence. Distinguishing between what can be said *of* and what can be said *about* something, and therefore what is true of and what is true about that thing, allows the creator of fiction to create a fictional entity to the exclusion of all others that cannot be accidently true. This is because, in the case of pure fictional entities, even where what can be said *of* the fiction and the actual, and therefore what is true *of* these things, is descriptively identical, what can be said *about* the fiction does not correspond to anything actual, and so cannot be true, even by coincidence.

The purpose of delineating the nature of pure fiction's existence was to provide a clearer understanding of fictional-x_i *qua* the target of our moral scrutiny. Now that we have established the nature of pure fiction, we are in a better position to determine what is required for a depiction of fictional immorality to be immoral and/or one's engagement with the depiction to be immoral. It is toward a more detailed examination of this matter that I now turn.

NOTES

1. Brown (2009) suggests that the fictitious murder may have been inspired by real-life events. https://www.theguardian.com/books/2009/apr/11/dickens-oliver-twist-eliza-grimwood-murder (accessed February 21, 2019). The inspiration, however, concerns the brutal nature of the attack, rather the circumstances leading up to the murder or even modus operandi of the attacker.

2. It does not hold unless one interprets the "conditions for existence" in a relativistic way, which, of course, defeats the universality of the claim.

3. One might argue that, in the case of private fantasy, the act of thinking is necessary and sufficient to bring into existence a fictional character. In reply, I would say,

first, that fantasy in and of itself does not satisfy the notion of depiction I am employing, here (as noted in chapter 1). Second, the act of thinking per se is not sufficient. After all, thinking about how hungry I am does not create a fictional entity. Perhaps, more importantly, however, even if thinking is necessary, it is not the fictional character that is doing the thinking but, rather, the author/creator (a point I shall return to). In other words, the fictional character does not itself think.

4. Ludlow (2006) makes a related argument when claiming that fictional entities do not suddenly become real; rather, they are always real. The idea that they are sometimes real, sometimes not, stems from the fact that fictional entities satisfy truth conditions only in certain context. I suggest that this is the case regarding that which exists. Fictional entities can be said not to exist sometimes but not others. Whether they do or not is therefore determined by the conditions of existence they are required to meet. Where it is said that fictional entities do continue to exist, they do so only insofar as they are required to satisfy certain conditions for existence and not others.

5. To reiterate an earlier point, one might claim that thinking is a necessary condition for the existence of fictional characters, even if it is not necessary for the fictional character to think. In a world devoid of cognition, there would be no fictional characters to perceive, recall, imagine. I will return to this point, shortly.

6. Once created, one is left to ponder *where* the fictional character is located (Brock, 2010). One might surmise that fictional entities are "located" in the abstract world created by the author and maintained by the literary community.

7. Arguably, even when referring to %#!, the example partially succeeds in communicating something meaningful because detectives do exist.

8. See also Davies (1996).

9. Abstract objects such as numbers or "justice" are subject to the fidelity constraint because we incorporate numbers and the notion of "justice" (and such like) into our everyday lives. The number 3 must be used, for example, in conjunction with the items one is counting (when there are three). Fiction, of course, is integrated into our lives, also, but not in a manner that must conform to the fidelity constraint. If one were confronted with a real-life situation and said out loud, "What would Sherlock Holmes do?" then either one is using reference to the fictional Holmes as a way of asking oneself, "What is the rational answer?" or, if one simply sought to step into the shoes of Holmes, as it were, for personal amusement and/or escapism, then, to do so, one would have to enter the fictional realm and thereby remove oneself temporarily from the real world and the fidelity constraint.

10. Currie accepts that if it was later revealed to me (in therapy, say) that I had witnessed the murder of Kitty Jane as a young child but repressed the memory, then that would no longer make the fiction *accidently* true.

11. The name, John Smith, as an abbreviated descriptor, is a nonrigid designator (Kripke, 1980) because it does not denote the same object in all possible worlds.

12. The same can be said of the following proposition: "He found her more beautiful than anyone else." This sentence could be equally descriptively true of fictional John Smith as actual John Smith, even though neither description, when said *about* one of them, is also about the other, or who the other finds the most beautiful.

Chapter 3

The Content of Fictional Immorality

INTRODUCTION

Depictions of violence in entertainment are as old as civilization itself. For example, *The Epic of Gilgamesh*—arguably the oldest recorded work of literature (Mesopotamia, circa 2000 BCE)—contains depictions of fictional violence, as does Homer's poem, *The Iliad* (circa eighth century BCE), along with numerous Elizabethan and Jacobian plays, including Shakespeare's. More recently, film, television, and video games have enabled audiences to experience a level of graphic violence, particularly in relation to fictional immorality, previously unknown (inter alia, films such as *Flowers of Flesh and Blood* from the Japanese Guinea Pig series and Fred Vogel's *August Underground*; TV series like *Game of Thrones* and *The Walking Dead*; and video games such as *Manhunt 1 & 2*, *God of War*, and *Hatred*). Recall, also, the words of the former South Australian attorney general, John Rau (from chapter 1), who, after watching the unlicensed *A Serbian Film*, referred to some of the scenes as so depraved that he was not prepared to describe them in detail.

Among other things, *A Serbian Film* depicts fictional acts of torture and sexual depravity, including necrophilia, rape, and child sexual abuse. If what is depicted had occurred for real, it would amount to a series of shocking, egregious acts; but, as an example of pure *fictional* immorality, what is depicted is not intended to represent actual events directed at persons alive or dead: it is pure fiction.

Given the account of pure fictional immorality presented in chapter 2, the question that concerns this chapter is: What are the conditions, if any, that need to be met in order to make depicting a purely fictional act of depravity, or indeed any purely fictional act of immorality, immoral?

In chapter 1, I presented three possible moral judgments (M1–M3) about fictional *content* based on the relationship between fictional-x_i and what fictional-x_i depicts (namely, x_i). These were:

M1: Because what is depicted by fictional-x_i is a depiction of immorality (x_i), *all* cases of fictional-x_i are immoral.

M2: Even though immorality (*qua* x_i) is depicted by fictional-x_i, this is not reason enough for *any* instance of fictional-x_i to be immoral.

M3: Whether fictional-x_i is immoral or not depends on the type of immorality (*qua* x_i) depicted by the fiction. Therefore, even though fictional-x_i always depicts immorality, only in *some* cases is the depiction itself, in virtue of what it depicts, immoral.

In this chapter, I intend to examine each of these judgments in turn. I will argue that only M2 should be upheld, even in the case of fictional immorality that depicts child sexual assault or includes characters that use racial (or other discriminatory) slurs.

ARE ALL CASES OF FICTIONAL IMMORALITY IMMORAL?

According to M1, the immoral nature of what is depicted in cases of fictional immorality is sufficient to render the depiction itself immoral. Under M1, it would be immoral to depict, as a fiction, *anything* that is immoral. This means that in addition to morally lambasting the more extreme depictions of immorality that feature in fictions such as *A Serbian Film* or Bret Easton Ellis' *American Psycho*—a position some (many) may find intuitively appealing—M1's all-encompassing moral condemnation would be directed toward the countless murder mysteries we regularly consume, alongside other violent crime dramas, as well as seemingly benign cases of fictional immorality, such as theft (e.g., films such as *The Sting*, *The Italian Job*) or infidelity (e.g., *Indecent Proposal*, *Closer*) or fraud (e.g., *Wall Street*, *Dirty Rotten Scoundrels*). These latter examples illustrate how M1 does not accord with our typical views on fictional immorality: for, typically, we do not hold that *all* depictions should be considered immoral. How should we respond to this? Should we align ourselves with convention and reject M1 for being overly stringent and draconian, or instead take our typical views as evidence of a moral failing on our part and embrace M1 as a measure of how we *ought* to treat all depictions of fictional immorality?

To endorse M1, we would have to accept the following:

- m_1: Fictional-x_i is immoral *if* x_i is immoral.

Against m_1, Di Muzio (2006) notes that "depictions of violence do not per se belong in the category of the morally objectionable only because many instances of real violence do" (p. 280).[1] The immorality of the Holocaust is therefore not *sufficient* to make depictions of a fictional genocide immoral (e.g., Orson Scott Card's *Ender's Game*). Neither is the immorality of infanticide *sufficient* to make Rubens' work illustrating the biblical account of the *Massacre of the Innocents* immoral; nor should the immorality of rape be *sufficient* to condemn, morally, any work depicting rape (e.g., the rape of the Sabine women from Roman mythology or, more recently, the 1988 and 2002 films *The Accused* and *Irreversible*). This is because there is no *logical* entailment between the immorality of x_i and fictional-x_i, just as there is no logical entailment between the immorality of ϕ and "other than ϕ" (i.e., it does not follow that if all ϕs are immoral then all "other than ϕs" are immoral). Moreover, even though fictional-x_i is more closely aligned to x_i than "other than ϕ" need be to ϕ, owing to the fact that fictional-x_i constitutes a fictional version of x_i, this does not create a *conceptual* entailment of the kind that occurs between, say, the concepts "biological entity" and "corporeality."

ON THE POSSIBILITY OF A CAUSAL CONNECTION

I am not claiming that the lack of logical and conceptual entailments negates the possibility that there are other reasons to judge fictional-x_i immoral; m_1 does not disallow this. It may be, for example, that the immorality of x_i is *causally* sufficient for the immorality of fictionally-x_i, such that the former causes the latter to be regarded as immoral. It is my contention, however, that any causal connection would have to be indirect and contingent.

To illustrate: It is not uncommon for people to find depictions of fictional immorality harrowing and, in some cases, even capable of causing a strong aversive reaction. According to Simon Callow, Charles Dickens' own rendition of the murder of Nancy by Bill Sikes was often delivered with such brutal realism that people in the audience fainted.[2] In such cases, one could argue that fictional-x_i (in virtue of being a depiction of that which is immoral) is immoral because of the harm (*qua* distress) it causes. If one were to argue this, however, at best, it makes fictional-x_i's immorality conditional on (i) the harm (*qua* distress) caused—that is, fictional-x_i is immoral *if* it causes harm (*qua* distress)—and (ii) the fact that the harm was caused by what the fiction depicts: that is, *if* the harm caused was the result of fictional-x_i depicting that which is immoral.

Fictional content can, of course, cause distress without depicting immorality, thereby negating the immorality of the thing depicted as a necessary condition. In a self-penned article published in *The Telegraph* in 2005,[3] author

Chuck Palahniuk describes how people regularly fainted during readings of his then-unpublished short story *Guts*. (At the time of the article, sixty-seven people were reported to have fainted, even though what Palahniuk depicts is not immoral, although it is horrific.[4]) But even when restricted to a sufficient condition, the relationship between the immorality of x_i, the fictional depiction of x_i (*qua* fictional-x_i), and the distress caused does not present the immorality of x_i as sufficient for the immorality of fictional-x_i; rather, it requires that fictional-x_i is immoral in virtue of depicting immorality *if* the depiction causes harm (*qua* distress). The immorality of fictional-x_i is not therefore based on its content per se (*qua* what it depicts) but is instead contingent on a particular consequence of the content: in this case, the amount of distress caused. Moreover, for such a position to apply to *all* depictions of fictional immorality (in accordance with m_1), S would have to be distressed by (inter alia) the depiction of Fagin teaching Oliver pickpocketing skills in the musical *Oliver!* or by Henry Gondorff (Paul Newman's character) teaching Hooker (played by Robert Redford) how to be a better confidence trickster in *The Sting*, which, in either case, seems unlikely.

The matter of harm will be taken up further in chapters 5 and 6 as it is not the focus of our discussion here. Moreover, as a means of rejecting m_1 (and therefore M1), it could be argued that whether fictional immorality is deemed immoral in virtue of what it depicts depends on meaning and context (e.g., whether the depiction is gratuitous, exploitative, perceived to be endorsing a particular immoral view, and so on) and therefore cannot (or should not) be determined by content alone. Again, these factors will be considered elsewhere (see chapter 4, for example). In the meantime, I would like to consider whether depictions of fictional child sexual abuse and slurs of a certain kind qualify as exceptions to the entailment I am otherwise rejecting, and therefore to the claim that the immorality of x_i is sufficient to render fictional-x_i immoral, at least in *these* cases. If true, such a claim lends support to M3, which proposes that only *certain* depictions of fictional immorality are immoral, irrespective of context. Endorsing this possibility means that we must reject M1, which I do, owing to its overly stringent nature. As a general claim, it is demonstrably false that the mere fact that a depiction is a fictional representation of something immoral renders this fact sufficient to judge the fiction immoral. Does the parable of the Good Samaritan, for example, contain an immoral depiction because it describes a violent attack on a traveler by unscrupulous persons? Surely not. In rejecting the ubiquity of M1, let us nevertheless examine the possibility that exceptions exist to the otherwise-rejected claim that fictional-x_i is immoral *if* x_i is immoral. Let us consider whether *virtual* pedophilia[5] and certain kinds of slur support M3: the view that even though not all examples of fictional-x_i are immoral *if* x_i is immoral, at least one is.

In the next section, I discuss U.S. legislation created in response to the increased availability and accessibility of virtual child pornography. I focus on U.S. legislation simply because it is the most discussed in the academic literature. In addition, I present an argument in favor of a necessary connection between the immorality of x_i *qua* child pornography and fictional-x_i *qua* virtual child pornography. It is, however, an argument I intend to challenge. Before continuing, it is also important to note that, *for now*, I will equate virtual pedophilia with virtual child pornography, treating them as equivalent. I do, however, recognize that they need not be the same, so much so that I plan to discuss a potential difference, and the moral implications of this, from the section "Child Pornography as Child Abuse" onwards.

VIRTUAL PEDOPHILIA AND THE SELECTIVE IMMORALITY OF FICTIONAL IMMORALITY

The 1996 Child Pornography Prevention Act (CPPA) was the first attempt by U.S. Congress to respond to the digital era by alluding (rather than making explicit reference) to the *virtual* sexual imagery of children within its definition of child pornography. The new definition sought to criminalize not only that which depicts actual sexual activity involving a minor (in the case of the United States, someone below eighteen years of age) but also that which *appears* to depict a minor engaging in sexual activity or *conveys the impression* that a minor is involved (Bird, 2011; Rogers, 2009; Russell, 2008). In 2002, however, a ruling by the U.S. Supreme Court (in the case of *Ashcroft v. Free Speech Coalition*; based on a 6–3 decision) directly challenged the CPPA, claiming that aspects of the legislation were overbroad and therefore unconstitutional, insofar as they prevented freedom of expression (Kosse, 2004; Mota, 2002). Thus, the U.S. Supreme Court ruled that while "it remains illegal to make, show or possess sexually explicit pictures of children . . . [there is] no compelling reason to prohibit the manufacture or exhibition of pictures which merely *appear* to be of children" (Levy, 2002, p. 319). It is important to make clear, though, that the 2002 ruling did not affect the continued prohibition of "morphed" images: namely, images of real children that had been integrated with some other image, or in some way altered, to create child pornography (Karnold, 2000).

In response to this ruling, in 2003, U.S. Congress introduced the PROTECT Act (which stands for *Prosecutional Remedies and Other Tools to End the Exploitation of Children Today*).[6] The PROTECT Act sought to clarify the overbroad nature of terms within the CPPA (like *appears to be* or *conveys the impression*) by seeking to prohibit virtual images that are *indistinguishable from* or *virtually indistinguishable from* actual images of children. The measure of whether a virtual image is indistinguishable from an actual image of a child (or virtually indistinguishable) is based on the extent to which an ordinary person is able to tell the difference between the two. The PROTECT Act does not therefore prohibit drawings, cartoons, sculptures, and paintings of child sexual activity per se, given that such imagery *is* distinguishable to the average person. For the same reason, it does not criminalize (inter alia) plays and films such as *Romeo and Juliet* or *Titanic* or *American Beauty*, which contain footage of adult performers appearing as minors engaged in sexual activity, thereby alleviating a previous criticism leveled at the original 1996 CPPA: that it was overly restrictive.[7]

The PROTECT Act (section 1466A) does, however, limit the permissibility of such representations where they are considered to be obscene or “hard-core” (Bird, 2011). In other words, regardless of their distinctiveness from any imagery of actual children and therefore regardless of the medium used (meaning that drawings, paintings, and so on are *included*), if a virtual image of a child involved in sexual activity or of a sexualized child is judged to be obscene, then it is deemed to be a form of child pornography, subject to prosecution under the law.

Thus far, I have described legislation pertaining to the legal status of virtual child pornography, at least in the United States; but what about arguments supporting its moral condemnation, specifically? Bartel (2012) claims that virtual pedophilia is morally objectionable insofar as child pornography is morally objectionable because virtual pedophilia *is* child pornography, given that it *necessarily* involves the *depiction* of sexual acts involving children. Recall that, at this stage, I am equating virtual pedophilia with virtual child pornography. Thus, for Bartel, virtual pedophilia *qua* virtual child pornography is immoral because child pornography is immoral.

If Bartel is correct, then we have an example of fictional-x_i (*qua* virtual pedophilia and hence virtual child pornography) that is immoral in virtue of its relationship to x_i (child pornography). Such a relationship makes the immorality of child pornography sufficient for the immorality of virtual child pornography and hence virtual pedophilia, just as m_1 attests (fictional-x_i is immoral *if* x_i is immoral). It is worth noting, however, that Bartel’s argument rests on the inclusion of virtual child pornography as a form of child pornography. As such, while m_1 is compatible with Bartel’s position, the structure of his argument is applicable to virtual pedophilia and not, for example, virtual murder. One could not claim with any degree of credibility that the virtual unlawful killing of someone is immoral because it is a form of murder.

In an attempt to support the assertion that virtual pedophilia amounts to child pornography, as it necessarily involves the *depiction* of sexual acts involving children (and let us allow that the depictions are not of actual children), Bartel presents us with the following hypothetical video game:

> Imagine a video game in which the gamer is allowed to voluntarily commit an act of virtual paedophilia and the act is graphically depicted. In such a case, the graphic depiction of a character—who is clearly depicted as an adult—engaging in sexual acts with another character—who is clearly depicted as a child—would count as an instance of child pornography. While these may be *virtual* instances of paedophilia, they are still *actual* instances of child pornography. (2012, p. 13; emphasis in original)

Here, we see Bartel classifying virtual pedophilia as a form of actual child pornography. Bartel's definition of pornography, on which his understanding of child pornography is grounded, is taken from Rea (2001, p. 134). According to Rea, an object acquires the ontological status of *pornography*—if (a) the object is put to pornographic use and (b) it is reasonable to believe that the object will be used as pornography, in accordance with point (a), by most of the audience for which it was produced. Condition (b) is important: for although something may be treated as pornography by an individual or even a group of people (in accordance with condition (a)), it is not pornography unless it is treated as such by the majority of the object's intended audience. Thus, although a nude image published in a naturist magazine may be treated as pornography by some, it should not be labeled "pornography" because (I assume) the majority of its intended audience do not treat it as such. Contrast this with an equivalent image published in, say, *Penthouse* or *Hustler* magazine.

In order for Bartel's example of virtual pedophilia within a video game to satisfy Rea's definition of what it is for something to *be* pornography, the depiction has to be put to pornographic use, and it has to be reasonable to believe that the majority of gamers who engage in the act of virtual pedophilia treat the enactment as a form of pornography, even if not all do. For now, let us accept that this is the case. I will, however, return to this point in the section "Non-pornographic Virtual Pedophilia." Importantly, then, when deciding whether the virtual enactment is *pornography*, the fact that it does not involve actual children is something of a moot point, for the point is this: where the function of the virtual enactment satisfies Rea's (and therefore Bartel's) definition of pornography, the virtual act of pedophilia is *pornography*.

For the sake of argument, let us accept that within a particular video game an act of virtual pedophilia occurs that satisfies Rea's definition of pornography (i.e., it is put to pornographic use by the *majority* of its intended audience). Under such circumstances, the following applies:

a) Virtual pedophilia is pornography.
b) x is an act of virtual pedophilia.
c) Therefore, x is pornography.

Notice how I refer to virtual pedophilia as pornography rather than *child* pornography. I do this, first, because Rea is interested in defining pornography in the absence of specific content but, also, I do so for a reason that will become apparent as we progress. Now, one might respond to my use of the term "pornography," rather than child pornography, not unreasonably, by pointing out that if virtual pedophilia is pornography, then it is pornography which necessarily involves the depiction of children (or at least one child) and that this depiction is intended to elicit sexual arousal from its audience (see Bartel & Cremaldi, 2018, p. 49, for more recent discussion on this point). Prima facie, pornography involving children is child pornography. Indeed, this is precisely what Bartel (2012) claims when he states that virtual pedophilia is child pornography because it necessarily involves the depiction of sexual acts involving *children*. This is not an unreasonable position to adopt. Nevertheless, I have the following, alternate, response in mind. I am willing to accept that child pornography typically involves the *depiction* of a sexual or sexualized act, but I wish to consider more carefully Bartel's claim that it necessarily involves the depiction of *children*. On the question of children, I am willing to accept that the depiction need not involve actual children in order for it to satisfy Rae's definition of *pornography* (i.e., it could involve virtual entities in the form of children which are intended to elicit sexual arousal), but, importantly, it is my contention that while "not involving *actual* children, only virtual ones" is not a barrier to the depiction being classified as pornography, it is a barrier to it being classified as *child* pornography. The reason for this is that I consider child pornography to be synonymous with child abuse.

CHILD PORNOGRAPHY AS CHILD ABUSE

Child pornography can be and often is a record of serious sexual assault on children (Adams, 2010; Edwards, 2000; Tate, 1992). The 2007 *Convention on the Protection of Children Against Sexual Exploitation and Sexual Abuse*[8] agrees, stating in no uncertain terms that child pornography *is* sexual abuse (Kalim, 2013). Likewise, Eneman et al. (2009) note how "it is inextricably harmful to children . . . [because the] production of child pornography *requires* a child to be abused" (p. 5; emphasis added). Similarly, Mal Shervill, assistant commissioner of the Western Australia Police, has this to say: "For every [pornographic] image they download there's a child somewhere in the world who's defenceless and without a choice, being abused and degraded, so it is a form of child abuse" (cited in Simpson, 2009, p. 255).

To my mind, child pornography and child abuse are *ontologically equivalent*. Given this, the following should apply:

d) Child pornography is child abuse.
e) *x* is child pornography.
f) Therefore, *x* is child abuse.

Even if the image is not of a child being sexually abused (i.e., a photograph of a naked child in a bath or even a fully clothed child in a park), this does not detract from the fact that images of children used for pornographic purposes are still abusive. Furthermore, superimposing the image of a child onto another image, so that it appears to be engaged in a sexual act, is exploitative: for although it is not a record of actual abuse, and in a sense is no more “true” than a painting (Williams, 2003), the child whose image is superimposed is still being exploited by the fact that the image purportedly shows them engaged in some form of sexual activity (Eneman et al., 2009).

Child pornography, in virtue of the fact that it involves actual children, is *sufficient* for child abuse to have occurred, and for it still to be occurring to the children represented. In the case of virtual pedophilia, of course, no actual children are involved in the depiction, and so no actual child is abused. This being the case, if we equate virtual pedophilia with child pornography, then we have a situation in which child pornography both does and does not involve the abuse of actual children. This would mean that, ontologically, child pornography both is and is not equivalent to child abuse.

Irrespective of current categorizations within legislation—in which virtual pedophilia is typically classified as child pornography—and given that my interest is in the ontological and moral status of virtual pedophilia and not its legal classification, it is my contention that the occurrence or not of child abuse should count as a pertinent ontological and, importantly, moral distinction, and therefore constitute a relevant means of moral discrimination in the cases we are discussing. Given this, we have a means of morally discriminating between child pornography and virtual pedophilia in virtue of the former’s ontological equivalence to child abuse and the latter’s lack thereof. It should not be difficult to accommodate this distinction into our discussion while accepting that virtual pedophilia is capable of satisfying Rea’s definition of pornography. Somewhat unremarkably, we simply distinguish between child pornography and virtual child pornography and equate virtual pedophilia (for now) with the latter and not the former; either that or we are forced to use a phrase like “non-abusive child pornography,” which does not seem appropriate, and in fact seems ripe for misinterpretation. In light of my proposal, consider the following:

(1) Child pornography is child abuse.
(2) Non-child pornography is not child abuse.
(3) Virtual child pornography is a form of non-child pornography.
(4) Given (2) and (3), virtual child pornography is not child abuse.

I accept that the claim "virtual child pornography is a form of non-child pornography" may seem like an odd thing to say; it does, however, make sense if one thinks of it as part of a broader classification of pornography said to be homogenous *only* insofar as none of it involves depictions of actual children, even if, in the case of virtual pornography, it depicts in a fictitious manner (inter alia) child sexual activities. With virtual child pornography, the depiction can be *of* child sexual abuse (insofar as that is what the image is meant to depict) without the depiction itself being a record of actual abuse.

To ease the awkwardness of the labeling I am using here (and its somewhat clunky fit), let us think of all forms of pornography, other than pornography involving actual children, as non-child pornography, and refer to this simply as pornography. If we do this, then any labeling or ontological confusion dissipates (at least when contrasting this much broader category with child pornography, specifically), as we can see:

(5) Child pornography is child abuse (it is a sufficient condition for abuse to have occurred or still be occurring).
(6) Pornography does not amount to child abuse.
(7) Virtual child pornography is a form of pornography and not a form of child pornography.
(8) Given (6) and (7), virtual child pornography is not child abuse.

Is pornography, including virtual child pornography, immoral? Adequately addressing this question would require a monograph of its own (although, see chapter 11 for some discussion on a relatively new form of pornography commonly referred to as *deepfake*). My argument in the sections "Virtual Pedophilia and the Selective Immorality of Fictional Immorality" and "Child Pornography as Child Abuse" is designed only to show that virtual pedophilia, as a form of virtual child pornography, is not equivalent to child pornography. As such, the immorality of child pornography is not sufficient to render virtual pedophilia (*qua* virtual child pornography) immoral. As I have said before, accepting this does not negate the possibility that a legitimate reason exists to denounce virtual pedophilia (*qua* virtual child pornography), even where the content is not deemed to be child pornography. I will therefore return to the issue of the morality of virtual pedophilia periodically throughout the monograph, including below.

NON-PORNOGRAPHIC VIRTUAL PEDOPHILIA

As a challenge to the possibility that virtual pedophilia is immoral if (and therefore because) child pornography is immoral, I have shown how the immorality of child pornography is not sufficient to render virtual pedophilia (*qua* virtual child pornography) immoral. In addition, it is also important to note that depictions of fictional child sexual abuse, and therefore depictions of fictional (including virtual) *pedophilia*, need not be classified as virtual child pornography if such depictions fail to satisfy Rea's definition of pornography. This fact pulls apart, even further, the relationship between child pornography (which is necessarily child abuse) and fictional depictions of pedophilia, some of which may not satisfy our accepted definition of pornography. To support this last claim, consider the oft-cited example of virtual pedophilia introduced by Luck (2009) when discussing the morality of video game content:

> Imagine you are playing a computer game, the object of which is to steal the Crown Jewels from the Tower of London. One way to achieve this goal is to seduce and sleep with a Beefeater's daughter, who just so happens to be 15. A player who commits this act of virtual paedophilia may do so, not because he enjoys the notion of having sex with a child, but because he wishes to complete the game. (2009, p. 34)

Or, more recently, Ali's (2015) example; again, discussed in the context of video game morality:

> In *Silent Hill 2* . . . the gamer controls a character who has murdered his own wife. The gamer controls this character as he uncovers the repressed truth about what he has done. Consider now the possibility of a *Silent Hill* game that takes on an equivalent scenario involving pedophilia . . . [I]t is not clear that a virtual pedophilic act in that game would be impermissible. (p. 272; emphasis in original)

Each of these examples involves the depiction of fictional pedophilia, but in a manner that does not conform to Rea's definition of pornography. In the case of sleeping with the Beefeater's underage daughter, according to Luck, this is done by the character (*qua* avatar) in the game as a means to an end—a way of stealing the Crown Jewels—and by the player of the game (who controls the avatar) for strategic purposes: to achieve a goal within the game. With Ali's fictitious version of *Silent Hill*, the depiction is used to facilitate story development. If either depiction is to be considered immoral, then it cannot be because they are examples of pornography, because neither is (at least insofar as they fail to satisfy the conditions set by Rea). As such, if these or similar fictional depictions of pedophilia are not examples of pornography (*qua* virtual child pornography), then they cannot be immoral *if* pornography is immoral. Not because pornography is not immoral (this has not been determined and will not be taken up in this monograph) but, to reiterate, because these examples of virtual pedophilia are not examples of pornography. Perhaps, instead, they are immoral simply because pedophilia is immoral. If so, then we find ourselves confronted with M1 again and forced to consider why we should accept *this reason* in the context of fictional depictions of *pedophilia* that are not considered pornographic but be unwilling to draw the same conclusion when talking about fictional murder or theft or infidelity or rape.[9]

Thus far, I have argued that virtual pedophilia is not a form of child pornography and is therefore not immoral in virtue of the fact that child pornography is immoral (contra Bartel, 2012). I have further argued that virtual pedophilia could satisfy Rea's conditions for pornography but have left unanswered the question of whether pornography is immoral. If one were to claim that pornographic virtual pedophilia—*qua* a form of non-child pornography—is immoral, then it has to be because all or at least some pornography (in addition to child pornography) is immoral *and* pornographic virtual pedophilia belongs to this category. Support for this claim needs to satisfy two conditions, however. First, an argument for why at least some pornography (other than child pornography) is immoral needs to be presented. Cases of pornography involving actual sexual violence or bestiality would lend strong support to such an argument, I contend (see, also, chapter 11). Second, even if such an argument were forthcoming, for pornographic *virtual* pedophilia to be placed within this category, it needs to be explained how a *fiction* could be capable of achieving the same nominal threshold of immorality as pornography involving sexual violence and bestiality. By saying this, I am not suggesting that the same explanation has to be found for each type of pornography but only that it has to be capable of reaching the same conclusion: that *this* is immoral.

Even if we accept that some cases of virtual pedophilia satisfy Rea's conditions for pornography (or at least could do so in principle), and even accept, for the sake of argument, that pornographic virtual pedophilia is immoral, examples of virtual or otherwise fictional pedophilia are nevertheless available that fail to meet Rea's conditions. In addition to Luck and Ali's examples, Nabokov's *Lolita* springs to mind, as does Jennifer Haley's 2013 play, *The Nether*.[10] One could also include the fictional depictions of the sexual grooming of a minor that occur in David Slade's 2005 film, *Hard Candy*. Each example illustrates how the immorality of pedophilia is not *sufficient* to make virtual/fictional pedophilia per se immoral. If it were sufficient then, to reiterate, an argument would need to be presented indicating why fictional-x_i is immoral *if* (and therefore because) x_i is immoral in the case of pedophilia/virtual pedophilia but not, say, in the case of murder/virtual murder or rape/virtual rape, and so on. In the absence of such an argument, virtual pedophilia per se is unable to support judgment M3 (I will return to the issue of pornographic virtual, or otherwise fictional, pedophilia in later chapters, particularly in relation to the issue of meaning [chapter 4], harm [chapters 5 and 6], and enjoyment [chapter 7]).

Actual pedophilia is abhorrent. This fact is nevertheless insufficient to justify the claim that depictions of fictional pedophilia are in and of themselves immoral. Before completely rejecting M3, however, there is one more example of immorality that I would like to consider. I do so in order to reject it, thereby enabling M3 to be rejected. The type of immorality I would like to consider is racial or otherwise discriminatory slurs.

SLURS AND THE SELECTIVE IMMORALITY OF FICTIONAL IMMORALITY

In a newspaper article, Sabur (2018) reports the decision of Minnesota schools to remove Harper Lee's *To Kill a Mockingbird* and Mark Twain's *The Adventures of Huckleberry Finn* from their curriculum on account of the fictions' use of racial slurs, particularly the N-word, and the possibility that they might offend and marginalize some of the students studying the texts.[11] Slurs are disparaging remarks often targeted at minority groups that imply stereotypical beliefs on the part of the utterer about the target group. Words in and of themselves are amoral but can be put to moral or immoral use depending on the intentions of the speaker. Some words, of course, are considered profane and are likely to offend or at least provoke a negative reaction, irrespective of context or intention (the C-word for vagina springs to mind). Yet, research supports the view that the use of profanity is often correlated with honesty (Feldman et al., 2017; Vingehoets et al., 2013); it acts as an indicator of how we truly feel about something (e.g., "You're a *god-damn* liar" or "fan-*f*cking*-tastic"). The use of profanity is also something we learn, in terms of when, where and with whom it is acceptable to be profane (Jay & Janschewitz, 2008). Of relevance to this discussion, however, is that profane words are sometimes used as, or in conjunction with, racial or other slurs (e.g., "Bitch" might form part of the more general profanity "Son of a bitch" or be directed toward a woman as a sexist slur, or even be used in conjunction with other profane words for emphasis: "You *f*cking bitch*). As such, despite the amorality of words, where the use of a word or phrase (hereafter, word) is *intended*, racially (or in some other illegitimate way), to *deride* and *discriminate*, then its *use* is immoral: for the act the word's use is designed to perpetrate (i.e., racially motivated derision and discrimination) is an immoral act.

In the novels mentioned above, characters are depicted using racial slurs as a form of racially motivated derision and discrimination. They are therefore depicted as acting immorally (at least by the standards of today), but the books themselves are not considered to be immoral. (I will have more to say on the relationship between an individual depiction and the worldview of the fictional narrative, taken as a whole, in the next chapter.) But what is the difference between depicting a fictional racial slur and an actual racial slur? After all, a racial slur expressed in a fiction looks and sounds exactly like a racial slur uttered anywhere else. Because of this, I may be offended by a word used by a character in a fiction, say, when addressing a member of a minority group, in much the same way as I would be offended if I heard the same exchange on the street. Despite the potential for offence, however, the mere use of the slur is not itself sufficient to make the utterance immoral, and this is so irrespective of whether the utterance is made by a fictional character or an actual person. At the same time, the fictional status of context in which the utterance occurs fails to negate the possibility that what is said is immoral.

To support this last set of claims, in a study by O'Dea et al. (2015), differences in offence taken at the use of the N-word (by a white person toward a black person) were reported by witnesses depending on whether the exchange was between friends compared to strangers. If we infer from this that the lessening of the offence taken correlates with the lessening of the moral outrage, then we have a case in which the use of the N-word (for example) is not necessarily judged to be immoral. It depends on who said what to whom and in which context. A view supported by O'Dea et al. when they state:

> Slurs are not always used in a derogatory manner . . . [Rather,] research on slurs has found support for the positive use of derogatory slurs among members of the group targeted by the slur. Members of stigmatized groups use appropriated slurs to enhance bonding and express affiliation amongst ingroup members . . . Research thus indicates that, in some situations, slurs are perceived to be less offensive and more socially acceptable. (2015, p. 155)

Similarly, Allen (2015) discusses the different uses of the N-word in Quentin Tarantino's film, *Pulp Fiction*. He notes how it is used as both a slur—that is, in a derogatory manner toward someone—but also as a means of conveying in-group affiliation. In this latter regard, consider the following exchange between the hired guns Jules and Vincent (taken from Allen, 2015, pp. 191–192):

Jules: You remember Antwan Rockamora? Halfblack, half-Samoan, usta call him Tony Rocky Horror.
Vincent: Yeah maybe, fat right?
Jules: I wouldn't go so far as to call the brother fat. He's got a weight problem. What's the nigger gonna do, he's Samoan.

Allen uses this extract to demonstrate Jules' use of the N-word as an in-group marker, toward someone he seems to have some liking for. It is not, then, meant to deride Antwan (a.k.a. Tony Rocky Horror).

What O'Dea et al.'s research and the *Pulp Fiction* example is intended to show is that intent and context are important. In actual exchanges, they help mitigate the likelihood of offense being taken, as well as the extent to which the slur is judged to be immoral. Intent and context also apply to the way the slur is *depicted* within fiction. Is it intended to be depicted as an in-group marker or as a derogatory comment? For the sake of argument, suppose it is intended as a derogatory comment, as occurs in the aforementioned novels by Harper Lee and Mark Twain (in order to reflect the period's language and attitude—of some at least—especially the characters portrayed within the fiction) and also in *Pulp Fiction* when, for example, one character refers to Vietnamese and Korean shopkeepers (who he is planning to rob) as gooks (Allen, 2015). The character's use of the term "gooks," in this instance, is intended as a racial slur; it conveys (we are meant to infer) the character's racist attitude. So, is the use of an *actual* racial slur, in order to *depict* an immoral attitude by a fictional character, toward other fictional characters, immoral?

As a way to address this question, I would like to contrast racial slurs with examples of fiction that have included actual footage of immorality. Before continuing, however, a point of clarification: I realize that the inclusion of *actual* footage means that we are no longer discussing pure fictional immorality. Nevertheless, I consider this brief departure to be acceptable given that we are discussing *actual* slurs within a fictional context.

van Ooijen (2011) finds the killing of animals (e.g., the beheading and dismembering of a giant turtle) in Ruggero Deodato's *Cannibal Holocaust* to be a morally flawed aspect of the film. Indeed, it is reported that the animal deaths are something Deodato regrets (He was found guilty of animal cruelty). *Cannibal Holocaust* is controversial for other reasons, too. Among them is the inclusion of footage of real-life executions, including children. *Cannibal Holocaust* is not the only film to include footage of actual immoral acts, of course. Oliver Stone's *JFK* (for example) includes the original footage of the assassination of President Kennedy, including repeated close ups of the fatal shot to the head.

The actual death of an animal is an unnecessary price to pay if one wishes simply to *depict* the death of an animal in a film (for many, this would be a clear example of animal cruelty and an immoral act for this reason). Other ways are available, especially given modern technology. Similarly, it is not necessary to include actual footage of immoral acts within a fiction. The reason for the inclusion of these examples, here, however, is for contrast: for while it is not necessary to include actual depictions of animal or human deaths in order to depict corresponding fictional versions of immoral acts, how does one depict a racial slur within a fiction without using an actual racial slur?

One way is to invent a slur (a neologism) and inform the audience or imply that this is what it is. An example of such a slur can be found in Ridley Scott's film *Blade Runner*, where the term "skin job" is used in a derogatory way to refer to replicants (e.g., "Christ, Deckard, you look almost as bad as that skin job you left on the sidewalk"). Such an option is unavailable, of course, if one wishes to retain a level of authenticity within the fiction regarding contemporary language or language employed in the past: again, as is the case with *To Kill a Mockingbird* and *The Adventures of Huckleberry Finn*.

So, in the absence of invented racial slurs of the kind found in *Blade Runner*, given that depictions of racial slurs are actual racial slurs, are such slurs evidence of an exception to M1 and therefore support for M3? In other words, in the case of racial slurs, is the use of such a slur within a fictional context—say, toward a fictional black man—immoral? Given the discussion above about intent and context, I do not believe so.

What makes a racial slur moral or immoral is the use to which it is put, and where this is to depict fictional immorality, then this use (*qua* intent), as well as the depiction itself, is not immoral, irrespective of any offense caused. However, should the depiction of fictional immorality be used to *endorse* an immoral worldview then we have a case of immoral fiction. Discussion on this issue is postponed, however, until the next chapter. In the meantime, let us turn our attention to the remaining moral judgment: M2.

IN DEFENSE OF M2

If it is not the case, as M1 would have us believe, that *all* fictional depictions of immorality are immoral in virtue of the immorality they depict, and it is not the case, despite M3's assertion to the contrary, that at least some or even just one example of fictional immorality is immoral (again, for the reason discussed), then does it follow from this that we should accept M2: the view that the immorality of x_i is not sufficient to render fictional-x_i immoral? In short, yes; but not simply because there is no alternative left. A more qualified and nuanced reason is required.

To be clear, all that M2 asserts is that the mere fact that fictional-x_i depicts that which is immoral is not sufficient to render the depiction immoral. It is saying that fictional immorality should not be considered immoral for *this* reason, and this fact applies to all cases. There may well be other reasons to render fictional-x_i immoral. Accepting M2 simply means that the immorality of what the fiction depicts is not (should not be) among them. It could be argued, however, with some justification, that examining the *content* of fictional-x_i is, in and of itself, a fairly ineffective task. Depictions of fictional immorality are not created in isolation. They are not like some fluke natural erosion of a rock face that seems to "depict" a face, or like the "image of Jesus" that appears on a slice of toast, or when viewing a cloud formation from a certain angle, and so on. Unlike each of these, fictional-x_i is intended to depict fictional immorality. It is intentionally created with this goal in mind. A pertinent question, then, is why did the creator of the fiction intend to create *this* fictional immorality? Or even this fictional immorality *in this way*? Considering these questions further will make a more productive contribution to our moral examination.

By way of a cursory example to preempt the discussion to come, suppose, as part of a review of *A Serbian Film*, or even as part of the classification board's report, S were to describe some of the scenes from the film that depict the gravest immorality (e.g., child sexual assault), perhaps including a "still-shot" to illustrate graphically whatever point S was trying to make. Under such circumstances, the reason for *this* depiction, in *this* context (e.g., S's review or report), would appear to be different to whatever reason the creator(s) of *A Serbian Film* had for depicting the fictional immorality so vividly to begin with. In each case (whether in the film or in the review/report), what is depicted may shock and upset the recipient (as touched on earlier with the example of the murder of Nancy by Bill Sikes, or in the case of racial slurs and virtual pedophilia), but whether this is reason enough to render the depiction immoral may depend on any number of other factors, yet to be examined, such as context, intention, meaning, consequence (intended or otherwise). Leaving all of these aside, for now, what I am rejecting is M1's general assertion that the mere fact that fictional-x_i depicts that which is immoral, irrespective of the immorality depicted, is reason enough to render it immoral. I also reject M3's more selective position, which states that moral condemnation should apply to depictions of certain immoral acts/events but not others. Given the absence of alternatives, logic dictates that M2 should be endorsed. Importantly, though, it should be endorsed for reasons other than its status as the only remaining option; it should be endorsed, I contend, because of the strength of evidence and argument which leads to the conclusion that, without exception, the mere fact that fictional-x_i depicts that which is immoral is not reason enough to render the depiction likewise immoral.

CONCLUSION

In conclusion, what I hope to have shown is that the morality of fictional-x_i cannot be determined by the morality of that which is depicted by fictional-x_i.

This does not mean that there are no other reasons to render fictional-x_i immoral; it just means that those hoping to render the fictional content per se immoral in virtue of what it represents will have to bite the bullet and accept that, in and of itself, it is not immoral. In the ensuing chapters, I will systematically explore what these other factors might be, starting with the *meaning* of the depiction.

NOTES

1. Di Muzio's comment likely includes depictions of historical as well as fictional violence. My focus, however, remains on depictions of fictional violence or other immoral acts/events.

2. See Callow's 2014 article: https://www.bl.uk/romantics-and-victorians/articles/dickens-the-performer (accessed April 19, 2019).

3. https://www.telegraph.co.uk/culture/books/3643352/67-people-fainted-as-I-read-my-horror-story.html (accessed April 19, 2019).

4. It is my contention that, typically, the majority of people today would not describe the acts depicted as immoral.

5. A point of clarification: The perpetrators of child sexual abuse are often referred to in the popular press as pedophiles. The clinical use of the term "pedophile" is, however, reserved for those who have a sexual interest in prepubescent children. Those with a sexual interest in pubescent and prepubescent children are known as a hebephiles (Neutze et al., 2011). Based on this clinical definition, someone can be convicted of violating child pornography law who is not a pedophile (because their sexual interest is in pubescent children). For convenience, however, the term "pedophile" will be adopted in a manner consistent with its more general, nonclinical, use.

6. In 2008, the PROTECT Act was upheld by the Supreme Court as constitutional.

7. Part of the 1996 definition of child pornography states that images are prohibited if they *appear* to depict an actual minor. Thus, age is determined by appearance. As Gillespie (2010) notes: "A child is to be taken as a child if it appears to be one" (p. 23). Jenkins (2001), however, notes a potential problem with the appearance criterion: "Some years ago, millions of people worldwide saw the film, Titanic, in which Kate Winslett plays a seventeen-year-old girl who has sex during the course of the story. Nobody was troubled by this incident, as the actress herself is well over the age of consent, but the film probably violated [U.S. child pornography laws] by simulating a sex act by someone presented . . . as a minor. Other recent films, such as Lolita and American Beauty, have faced similar dangers" (p. 220; cited in Gillespie, 2010, p. 23). The reason why nobody was troubled by the Kate Winslett sex scene in *Titanic*, Gillespie (2010) conjectures (with some justification), owes to the fact that Kate Winslett did not appear to be seventeen years of age; rather, she was presented as being this age in the film. U.S. child pornography law is concerned with how old a child *appears* to be not with how old they are presented as being. Nevertheless, the PROTECT Act, among other things, tried to make the definition of "appearance" clearer.

8. Available at http://www.coe.int/t/dg3/children/1in5/Source/Lanzarote%20Convention_EN.pdf (accessed August 9, 2016).

9. The issue raised here is similar to something Luck (2009) refers to as the gamer's dilemma.

10. https://www.nytimes.com/2015/02/25/theater/review-jennifer-haleys-the-nether-explores-the-dark-side-of-the-web.html (accessed April 29, 2019).

11. See also Olulode (2015).

Chapter 4

The Meaning of Fictional Immorality

INTRODUCTION

In the previous chapter I stated that fictional content is rarely, if ever, presented or understood in isolation. When engaged with a fiction (and, at this point, I am still restricting this to fictional-x_i), aside from recognizing that the content is fictitious, the audience will likely extract meaning from it, or at least attempt to do so—say, by seeking the moral of the story—based on an expectation that there is one to be had. As part of the process of comprehension, Booth (1988) notes how we typically distinguish between two components of a fictional narrative: *nonce beliefs* and *fixed norms*.

Nonce beliefs are those we are required to hold for the duration of the fiction (e.g., zombies walk the Earth in ever-increasing numbers). They may well constitute a nonrealistic fiction, insofar as the fictional reality depicted does not represent current norms (i.e., we are not in the midst of a zombie apocalypse), nor are the make-believe truths that constitute nonce beliefs intended to transcend the fictional domain: that is, be exported from the fictional world to the nonfictional. Importantly, this same constraint on nonce beliefs applies equally to fiction that is nonrealistic in virtue of inverting *moral* norms. Liao (2013) illustrates this with reference to Joseph Heller's satirical novel *Catch-22*. According to Liao, *Catch-22* invites us to fictively imagine a world in which moral norms are, for the most part, or in significant ways, inverted and therefore, by our standard, immoral. Ultimately, for Liao, this is done to enable the audience to reflect on the absurdity and immorality of war and all that goes with it. Thus, the fiction employs make-believe immorality as a means of exporting from the fiction a moral position, or at least to invite moral reflection outside of the fictional world (a point I shall return to later in this chapter and also discuss in relation to the phenomenon of imaginative resistance in chapter 8).

Also contained within the fiction are fixed norms. These may be physical norms (i.e., in *Catch-22*, the airplanes do not defy the laws of physics) but, equally, may reflect moral norms such as being honorable, treating people with respect, good triumphing over evil, and so on (these are more consistently inverted in *Catch-22*). Nonrealistic fiction may violate physical laws while maintaining moral norms or vice versa, or indeed violate both. To illustrate: even though Darth Vader is understood to be part of the fiction (requiring a nonce belief), and to inhabit a world that cannot be accounted for by the laws of physics, at least as we understand them, the estranged relationship he has with his son, Luke Skywalker, might nevertheless be said to have a level of authenticity that transcends the fictional domain (Malliet, 2006).

Fixed moral norms provide the backdrop against which we are to judge the exploits of the protagonist (and others). They define the moral of the story, and hence what we are expected to take away from the fiction. Fixed norms are therefore meant to possess a transcendent quality that, in the case of realistic fiction, is either avowed or implied (realistic fiction is fiction that maintains moral, psychological, and physical norms). Again, Liao (2013) illustrates this with an example. In Jane Austin's *Pride and Prejudice*, it is never expressly stated that pride is a vice; rather, it is simply assumed (but also required for an authentic reading) that we would import this moral norm from our world into the fiction (see also discussion on Walton's notion of *work world* in chapter 7). As such, if aspects of the fictional world are not delineated, then, according to Ryan's (1991) *principle of minimal departure*, we fill in the blanks by extrapolating from what we do know about that world or by importing some understanding from the actual world.

Importantly, there is a symmetry to the fixed norms of realistic fiction. In *Pride and Prejudice*, we assume that pride is a vice (importing this assumption where necessary) and expect the fiction to convey this assumption by way of its moral message. Should nonrealistic fiction invert this norm and depict vice as a virtue, then it is in contrast to this fixed norm that we understand our invitation to engage with the moral inversion. The ultimate purpose for this engagement, however, and its moral implications, will be discussed in the sections to come.

Even in cases of fiction that does not have a spoken or written narrative, the distinction between fixed norms and nonce beliefs can be made, at least where one is able to comprehend the meaning—*qua* nonlinguistic narrative and/or the moral—of the imagery (e.g., Fritz Lang's silent film *Metropolis* and Pieter Bruegel the Elder's collection of paintings entitled *The Seven Deadly Sins*). One may not agree with or even like the stance presented within a particular fiction—whether in a book or film, play or TV show, poem, painting or song lyric, and so on—but one expects there to be one (Nussbaum, 1992). Should a stance be missing, then this can be interpreted as a norm violation and therefore as an exception that proves the rule. Thus, Nolan (2019) dismisses the idea that a novel like *Lolita* is little more than an instruction manual for pedophiles. Instead, it is intended to present a point of view, or perhaps a number of conflicting viewpoints, for the audience to digest, wrestle with, and ultimately judge. In the case of *Lolita*, despite depicting immorality, Nolan (2019), along with many others, holds that the book itself—its meaning and intention, and therefore what it depicts as a means of conveying this—is not immoral, because the author's meaning and intention are not immoral.

The aim of this chapter is to examine the role meaning plays in our decision making about the morality of fictional-x_i. In particular, I focus on the extent to which depictions of fictional immorality are congruent with a fiction's endorsement of an immoral worldview. Where individual depictions are congruent, it is my contention that these should be considered immoral. I recognize, however, that upholding such a claim, even in principle, is easier to do than applying it successfully to specific cases of fictional immorality: for it may not always be apparent, in a definitive sense, that a fiction *is* endorsing an immoral worldview and, therefore, that a depiction of fictional immorality is congruent with the alleged immoral worldview. Because of this, I consider an alternative position based on a *reasonable interpretation* of the meaning of the fictional content, taken as a whole, and whether, in accordance with this reasonable interpretation, one is justified in claiming that individual depictions of fictional-x_i, congruent with the immoral worldview, are themselves immoral or whether, instead, they are, minimally, morally insensitive. I also discuss moral and aesthetic edification (or perhaps "appreciation" is a more apt term, in the latter case) as a means of justifying depictions of fictional immorality, even where the meaning of the depiction within the fiction, taken as a whole, is ambiguous and potentially morally insensitive.

THE IMMORALITY OF ENDORSING AN IMMORAL WORLDVIEW

Recall from chapter 1 the following claim about a fiction that endorses an immoral worldview (IWV):

> IWV Where a fiction, taken as a whole, endorses an immoral worldview, then any depiction of immorality from within the narrative, congruent with that worldview, is immoral.

Nabokov's *Lolita* depicts an immoral worldview: namely, the worldview of the protagonist, Humbert Humbert, who is infatuated with "Lolita," a minor; but, importantly, although the fiction depicts this worldview, it does not endorse it. In fact, according to McGinn (1997), what Nabokov reveals is a kind of aesthetic paradox. Through Humbert Humbert, "spiritual repulsiveness expresses itself in the most exquisite and cultivated prose" (p. 110); and while this may unsettle us, aesthetically and morally, the individual depictions of immorality congruent with and constitutive of the immoral worldview depicted within *Lolita* (e.g., "I knew I had fallen in love with Lolita forever, but I also knew she would not be forever Lolita") fail to satisfy IVW. In a similar vein:

> One may be morally repulsed by the opening scenes of *Gone with the Wind* (1939), which show slaves happily at work in the fields of an ante-bellum plantation, bantering good-naturedly about which slave has the authority to tell the others that it's "quitting time." One may be similarly repulsed by the initial on-screen narrative, which extols the traditions of the "old South" as one of the high points of civilization, which were unfortunately destroyed during the Civil War. But there is ample reason to deny, based on later sequences in the film, that the film itself should be interpreted as endorsing slavery. After the war, for example, Ashley Wilkes—who, with his wife, constitutes the moral compass of the film—explains to Scarlett that he was going to free the slaves at Twelve Oaks, his family's plantation, once his father died. (Feagin, 2010, p. 28)

In contrast to these examples, recall from chapter 1 the video game *Ethnic Cleansing*. This fiction is the product of the white supremacist organization *National Alliance*, and it contains depictions of fictional violence directed toward various minority groups.[1] If the video game's content, taken as a whole, *is* intended to endorse and promote the organization's immoral worldview, then where the numerous individual depictions of violence directed at minorities presented within the game—for no other reason than they are from a minority group, or perhaps with the added incentive that they are stereotypically portrayed as acting unlawfully[2]—are congruent with the immoral worldview being endorsed, then each depiction satisfies IWV.

To endorse something is to declare, publicly, one's approval of and/or support for that thing. Public support and/or approval is difficult to give unintentionally, although I accept that one may do one, publicly, without expressly doing the other. What I mean by this is that while it may be that one publicly expresses both approval and support for ϕ, it could also be that one simply *expresses* approval, which is taken (not unreasonably) to be a form of tacit support; or, if one were simply to express support, as tacit approval. Either way, if a fiction *is* endorsing an immoral worldview—say the view that nonwhites are inferior beings and therefore not subject to the same status, rights, and moral considerations as whites—then the fiction, through its depictions taken as a whole, is approving of and/or supporting (likely both) the immoral worldview depicted, and arguably *promoting* this worldview to those exposed to the fiction. It is not simply *depicting*, in a fictional context, racial discrimination; it is using the fiction—the supposed make-believe—as a *means* or *vehicle* to express and, importantly, commend actual beliefs to its audience: beliefs that are immoral. In essence, the fiction is promoting an immoral worldview by conveying an immoral message through the fictional narrative, to the effect that what the fiction depicts—in this case, white supremacy and racial discrimination—has a veracity beyond the fictional realm.

The immorality depicted is therefore positioned as a fixed norm (to borrow Booth's term), at least from a white supremacist's perspective, that is intended to transcend the fictional domain and persuade its audience not simply to engage with the make-believe truth of fictional-x_i but, instead, support what it depicts by coming to *believe in* the truth of white supremacy as a justification for racial discrimination (see Liao, 2013). In such circumstances, while the fictional content is not immoral per se (as discussed in chapter 3), the *message* conveyed through the use of the fictional content is immoral. It is immoral not simply because of any potential to cause harm, such as increasing the likelihood of antisocial behavior toward minorities (although I will have more to say on this in chapter 5); rather, the message is immoral, first and foremost, because certain beliefs that are being endorsed through the message—in this case, the ideology of white supremacy—are immoral. The belief that white people are on average superior to any and all nonwhites is not, in and of itself, immoral, in the same way that the belief that the red team is superior to the blue team is not. In the former case, the belief lacks empirical support (based on conventional measures of "superior") and is false because of this. This erroneous belief can and often does ground other beliefs that are immoral, however, such as the belief that *how one is treated should depend on one's ethnicity*. This belief is immoral, even if it is never acted on, not only because of how it proposes we *ought* to behave—in an unjustifiably discriminatory manner (making the belief wrong in and of itself)—but because holding the belief, as I intend to show in chapter 5, makes people worse off in a way that is harmful.

It is also worth noting that certain fictions may inadvertently or tacitly endorse an immoral worldview, at least by the moral standards of future audiences (or even past audiences). The original 1960s *Star Trek* series, for example, has been rightly praised for many of its progressive social attitudes, including one of the first interracial kisses to be aired on U.S. television. Nevertheless, by today's standard, the sexism evident at times makes for uncomfortable viewing. Importantly, though, while sexist attitudes toward women were being challenged vociferously during the 1960s and beyond (reflected in films like *Rachel, Rachel*: directed by Paul Newman, 1968), it would be misguided, I feel, to claim that *Star Trek* was inadvertently endorsing something that, at the time, was accepted by all or even the majority as immoral. The same might be said of early James Bond films. This fact should not undermine the veracity of IWV, however. IWV is intended as a contemporary measure; it reflects today's dominant, even if not universally accepted, moral standard. As such, one should take it to mean the following: where a fiction, taken as a whole, endorses an immoral worldview, by the recognized standard of the day (see chapter 14 for detailed discussion on what this involves), then any depiction of immorality from within the narrative, congruent with that worldview, is immoral.[3]

AMBIGUOUS CONTENT: ENDORSING OR NOT ENDORSING A PARTICULAR VIEW?

My comments about *Ethnic Cleansing* and the worldview it promotes are, of course, conditional on the fictional narrative (and therefore the creator(s) of the fiction) actually endorsing an immoral worldview. In the absence of confirmation of this, however, given what is known about the National Alliance's white supremacist beliefs and the nature of the fictional violence depicted within the video game, such a conclusion nevertheless amounts to a *reasonable interpretation* of the gaming narrative taken as a whole. There is, of course, a difference between claiming that fictional-x_i is immoral because it aligns, in some categorical sense, with a fictional narrative that, taken as a whole, endorses an immoral worldview and the more modest, but perhaps more realistic, claim that fictional-x_i is immoral because it is congruent with a fictional narrative that one *reasonably interprets* in this way or that one could envisage being interpreted in this way by the majority of those who were presented with the fiction. This difference (between what *is* the case and what is reasonably *believed* to be the case) is highlighted below when one compares statements 1–4 with (a)–(d).

1. Fictional-x_i is immoral *if* the fictional narrative taken as a whole endorses an immoral worldview and fictional-x_i is congruent with that worldview (ontological claim).
2. The fictional narrative taken as a whole endorses an immoral worldview (ontological claim).
3. Fictional-x_i is congruent with the worldview endorsed by the fictional narrative (ontological claim).
4. Therefore, fictional-x_i is immoral (ontological claim).

In contrast:

(a) Fictional-x_i is immoral *if* the fictional narrative taken as a whole endorses an immoral worldview and fictional-x_i is congruent with that worldview (ontological claim).
(b) (Given S's reasonable interpretation of the fictional narrative) S believes that the fictional narrative taken as a whole endorses an immoral worldview or envisages (based on a belief about their likely interpretation) just such an interpretation from the majority of those who are presented with the fiction (epistemic claim).
(c) Fictional-x_i is congruent with the immoral worldview that (in accordance with (b)) S reasonably believes in being endorsed by the fictional narrative taken as a whole or would be interpreted this way by the majority of the audience (ontological claim).[4]
(d) Therefore, fictional-x_i is immoral (ontological claim).

With statements 1–4, the ontological conclusion arrived at in 4 is valid because it is based on ontological premises. Given 1–3, 4 must follow. In contrast, the conclusion in (d) does not follow, necessarily, from premises (a)–(c) because premise (b) describes a particular epistemic relationship, and an ontological conclusion cannot be deduced from an epistemic premise. To illustrate:

- S believes that she is meeting Robert Galbraith (author of the Cormoran Strike novels) for high tea (epistemic claim).
- S does not believe that she is meeting J.K. Rowling (author of the Harry Potter novels) for high tea (epistemic claim).
- Therefore, Robert Galbraith and J.K. Rowling are not the same person (ontological claim).

J.K. Rowling wrote the Cormoran Strike novels under the pseudonym Robert Galbraith. Therefore, J.K. Rowling and Robert Galbraith are the same person. The ontological claim that they are not is false. But even if it were true, it could not be deduced from the epistemic premises that precede it. In other words, the ontological status of J.K. Rowling and Robert Galbraith—that they are the same person—cannot be deduced from statements about what S believes and does not believe about them. Given this, if one accepts the truth of premise (a)—that fictional-x_i is immoral *if* the fictional narrative taken as a whole endorses an immoral worldview and fictional-x_i is congruent with that worldview—then the extent to which one is prepared to defend the conclusion in (d), given that it cannot be deduced from (a)–(c), depends on how much one can justify one's interpretation and subsequent belief that the fictional narrative (taken as a whole) endorses an immoral worldview or would be interpreted this way (for good reason) by the audience (as described in b).

To illustrate the potential difficulty with justifying premise (b), consider differences in the respective views of Ostritsch and Goerger when discussing the video game *Grand Theft Auto V* (hereafter, GTA). About GTA, Ostritsch (2017) has this to say:

> [It] is the representation of an immoral world. However, it would be wrong to infer from this that the game is also the endorsement of such an immoral world. Rather, anyone who has actually played the game will attest to its dominant satirical character. . . . [T]he world of *GTA V* can easily be identified as a satirically exaggerated version of our world, the real one. But as satire *GTA V* does not endorse what it portrays, rather it ridicules it. (pp. 123–124)

In contrast, Goerger (2017) holds that “there are good reasons for thinking that *GTA* is a morally problematic game” (p. 102). Goerger attempts to justify this claim as follows:

> Players are, essentially, being entertained by the misery of others and are thus disrespecting objects of value. . . . When . . . these considerations are brought together, they provide ample reason to think that deriving amusement from *GTA* reveals a defect of character. The game’s values are either not aligned with what merits respect or they fail to manifest themselves. (*ibid.*)

Both authors express a different view about the content of GTA: about the values espoused and therefore the morality of the gameplay. Which view, if any, is correct? Uncertainty over the answer to this question highlights a problem with IWV: namely, the ease with which it can be applied effectively to real-life examples. My own “reasonable interpretation” of the gameplay of GTA is different again: for even if we are to accept, as Goerger would have us believe, that playing GTA as the game designers intended, or at least permit, risks or even requires us to value incorrectly the experiences of certain people—for example, victims of homicide or rape, or minority groups or other *objects of value*—this is not necessarily the same as (*qua* correlated with) *endorsing* their inferior moral status. There is a difference, I would argue, between trivializing the experiences of a victim of immorality (i.e., someone murdered or sexually assaulted or subject to racial discrimination) and endorsing the view that such immorality is acceptable. To illustrate: If I endorse murder and commend others to do the same, at least in relation to certain minority groups, then I cannot at the same time hold (sincerely) that *all* murder is morally wrong (and, by this, I do not mean that I fail to recognize that others consider it to be wrong). If, on the other hand, I am guilty of trivializing murder, even in what might seem to be an amusing way—that is, imagine I created what is in fact a genuine, although withdrawn, advertisement for liquid soap in which a blooded arm is depicted in the foreground of a murder victim prostrate on the floor with a knife in his chest, accompanied by an image of the product alongside the caption "when ordinary soap just won't do"—then this is something I could have done without intentionally seeking to endorse the merits of murder.

It therefore seems reasonable to hold that, even if one fails to recognize and subsequently accept the satire some (e.g., Ostritsch) say is indicative of GTA (i.e., directed at society's Janus-faced attitude toward immorality, rather than established moral values), the alternative is not necessarily that GTA is endorsing immorality, intentionally or otherwise, as I hope the example of the advertisement for liquid soap illustrates. One could proffer the view that GTA is trivializing various forms of immorality (whether for satirical or other reasons) rather than endorsing them. Claiming this is not to downplay the need to examine the morality of trivializing immorality (I will return to this matter in chapter 9, when considering the difference between immorality and poor taste); rather, it is simply to defend the argument that, in supporting IWV, much rests on the veracity of the assertion that a particular fictional narrative *is* endorsing an immoral worldview. It would be difficult to determine precisely how many video games like GTA, or indeed any other depiction of fictional immorality, are actually endorsing immorality, rather than just *believed* to be so doing, however reasonable this belief may be.

Despite these problems, it is still important to support IWV: to make clear that what it asserts is, *in principle*, morally sound (Bartel & Cremaldi, 2018; Cooke, 2014). Thus, *if* the fictional narrative taken as a whole endorses an immoral worldview and fictional-x_i, in virtue of being a part of that fictional narrative, is congruent with that worldview, then fictional-x_i is immoral. This principle should hold irrespective of whether fictional-x_i depicts stealing, stalking, or stabbing, or any other kind of immoral act, or how well the message is presented in terms of aesthetic quality (e.g., Leni Riefenstahl's 1935 pro-Nazi film *Triumph of the Will*). Having said that, while it is important to champion IVW's moral stance, in practical terms, it is unlikely to be widely enforceable. Whether a fictional narrative, taken as a whole, *is* endorsing a particular immoral worldview may well be difficult to establish; it is, after all, unlikely to be something that is publicly expressed in an unambiguous form. So, even if it is the case, unless it is *known* and can therefore be shown to be the case, that a particular fiction is endorsing an immoral worldview, it is far more likely that different interpretations will be proffered, as the GTA example illustrates. In light of this very real difficulty, perhaps an amended version of IWV is required.

> $\mathrm{IWV}_{\mathrm{revised}}$ Where a reasonable interpretation of a fiction, taken as a whole, cannot establish whether the fiction is merely depicting rather than endorsing an immoral worldview, then any depiction of immorality from within the narrative congruent with that worldview is immoral.[5]

$\mathrm{IWV}_{\mathrm{revised}}$ takes its moral force from IWV, insofar as the former has moral veracity *only if* the latter has. By this, I mean that the moral legitimacy of IWV is necessary for the moral legitimacy of $\mathrm{IWV}_{\mathrm{revised}}$. What $\mathrm{IWV}_{\mathrm{revised}}$ attempts to do is provide the means by which the moral principle expressed by IWV can be practically applied, although not infallibly so.[6] $\mathrm{IWV}_{\mathrm{revised}}$ therefore errs on the side of caution. Thus, even where it is not known that ϕ is endorsing immorality, $\mathrm{IWV}_{\mathrm{revised}}$ holds that it is better to be safe than sorry *if* a reasonable interpretation of the fiction cannot establish that it is not. Such an amendment raises the following questions, however:

1. Is $\mathrm{IWV}_{\mathrm{revised}}$ morally justified?
2. If $\mathrm{IWV}_{\mathrm{revised}}$ is morally justified, then what could nevertheless justify the presentation of a fiction that (in all likelihood) satisfies $\mathrm{IWV}_{\mathrm{revised}}$?

IS $\mathrm{IWV}_{\mathrm{REVISED}}$ MORALLY JUSTIFIED?

$\mathrm{IWV}_{\mathrm{revised}}$ permits us to condemn, morally, a fictional narrative, along with congruent examples of fiction-x_i from within the fiction, even when it is not known that the fictional narrative is endorsing an immoral worldview. The moral condemnation stems, instead, from a reasonable interpretation of the fiction, taken as a whole. To determine whether $\mathrm{IWV}_{\mathrm{revised}}$ is justified, consider Powers (2003), who argues that the fictional content of video games, even though they are intended as *games*, is nevertheless capable of conveying a morally meaningful message (as, of course, are more traditional fictions such as literature, plays, songs, and more recently, of course, film and TV). Powers' point is that when one virtually enacts rape (for example) within the context of a video *game*, during one's attempt at playing the part of a rapist, the fiction one engages with as part of this play is nevertheless *in danger* of conveying—insofar as it might reasonably be misconstrued as conveying—the socially significant message "rape is okay," even if this is not one's intention while engaged with the fiction, nor the intention of the video game's creator. Equally, the fiction could be misconstrued as supporting historical positions on the subjugation and exploitation of women or, in the case of a video game like *Ethnic Cleansing* (but where one does not intend to endorse white supremacy), the dehumanization of other groups. Accepting that fiction is capable of conveying meaning—that is, socially significant expression—does not, however, provide guidance on when a fiction *is* endorsing an immoral worldview, but it may provide a means of justifying IWV$_{\text{revised}}$'s stance to err on the side of caution, given the potential significance of some misunderstandings in some contexts.

Patridge (2011), for example, talks about how certain video games have *incorrigible social meaning* (an accusation I am confident Patridge would agree translates to other forms of fiction). By this, she means that some content may depict something that has deep-rooted (actual) social meaning, even if only localized to a particular society, which may be deemed offensive to certain members of that society. She illustrates this with a fictitious example of a video game that depicts an African American navigating their way through a field of watermelons. The association of an African American with watermelons (and similar imagery), we are told, has "been used as a mechanism to insult and dehumanize African-Americans, and to bind racist Americans together through the practice of telling racially demeaning jokes" (p. 308).

The virtual character who walks through the watermelon field in Patridge's example is fictional but nevertheless represents a race of people who have been associated in U.S. history with watermelons in a manner that is now held by many in the United States to be morally offensive. Patridge recognizes that the association between this image and racism is a contingent fact and localized to U.S. society.[7] Different racist imagery may be found in different societies, each association being similarly contingent. Nevertheless, Patridge's point is that the *epistemic flexibility* we possess to create fictions that may be more or less loosely based on real-life contingent associations (objects/events), and therefore the extent to which we are willing in the pursuit of these fictions to suspend our understanding of their associations, ought to be constrained in relation to their potential incorrigible social meaning. Sometimes, she argues, we should reject fictional depictions if what they depict are associations that still have morally offensive undertones.

The virtual enactments (and other depictions) targeted by Patridge are those that were once held to be something of a social norm (e.g., institutionalized racism, at least within the United States) but that are no longer viewed in the same way. What she seems less concerned with are actual morally/legally prohibited actions that have never been a social norm. This is suggested when she states: "Consider, for example, the game *Mafia Wars*. The fact that we may play this game seems to say nothing at all by itself about our attitude toward organized crime" (2011, p. 307). Organized crime, as far as I am aware, has never been hailed as an acceptable social norm in the United States. Therefore, what I take Patridge to be saying is that if we play a game that features organized crime, this is not necessarily a sign of our endorsement of organized crime. On the other hand, "as morally challenging representational content begins to reflect our actual, shared history of systematic moral violations like gender and racial oppression, this serves to limit the meaning of such imagery" (*ibid.*, p. 310).

For Patridge, then, if the content of a video game—or, again, I would say any *depiction* of fictional immorality—reflects a previous endemic moral violation, then this raises legitimate concerns over the extent to which the content can be said to represent and therefore mean something other than the continued expression of the no-longer-tolerated but previously endemic violation. Moreover, although the creation or even the playing of such a game does not necessitate one's endorsement (support/approval) of what the fiction depicts, it does leave the fiction susceptible to $\text{IWV}_{\text{revised}}$, at least within the context of the society for which the depiction has incorrigible social meaning.

The idea of incorrigible social meaning provides a means of explaining the unease many of us feel toward certain depictions of fictional immorality, such as discrimination and violence toward minorities, or sexual violence toward women. When talking about video games that depict rape, Patridge argues that virtual rape has incorrigible social meaning because of the "global history and current reality of women's oppression" (*ibid.*, p. 312).[8] Patridge is, however, willing to concede, or at least suggests, that it is perhaps only in a world without gender oppression or other forms of discrimination that such fictional depictions are likely to lose their incorrigible social meaning. Perhaps, equally, one might claim that only in a world without historic and even current gender oppression and racial discrimination would fictional depictions of such acts be less likely to be interpreted as endorsing these respective immoral worldviews, and therefore be less likely to fall foul of $\text{IWV}_{\text{revised}}$.

Summarizing thus far: There may be nothing within the content of the fiction per se that necessitates the message "rape is okay," even if the fiction were to depict something as graphic and immoral as an alternate literary reality in which it was revealed that Charles Dickens' character, Jacob Marley, engaged in non-consenting sexual intercourse with an employee in front of a crowd of approving onlookers. Nor does one necessarily have to endorse the fictional narrative in which this is seen as acceptable to engage with the fiction; otherwise, an audience attending a performance of, say, Shakespeare's *Titus Andronicus* would pretty much have to be of the mindset "anything goes" to engage with that particular fiction. Nevertheless, in the case of certain fictional content, in the context of certain audiences, based on a contingent historical relationship between the immorality depicted by the fiction and the group/society the audience belongs to, there is a case to be made for erring on the side of caution when it comes to the possibility of fictions endorsing immoral worldviews. Or at least for mitigating circumstances to be required where there is ambiguity within the fictional narrative (see section

"Mitigation in the Form of Moral Edification"). That said, I am skeptical about whether it is possible to enforce a claim of immorality based solely on $\mathrm{IWV}_{\mathrm{revised}}$'s "reasonable interpretation" requirement, given the fact that it is based on *interpretation*, no matter how reasonable this may appear to be. I do, however, accept that it may be reasonable to charge fictions that *seem* to be endorsing an immoral worldview but where this is open to doubt and is therefore equivocal, minimally, with moral insensitivity.

MORAL INSENSITIVITY

Moral sensitivity refers to one's ability to recognize and *appreciate* the feelings and needs of others during one's interpersonal interactions or other behaviors (Christen & Katsarov, 2016; Lovett & Jordan, 2010; May, 1992; Smilansky, 1996). To be morally *insensitive* is therefore to fail to exercise adequately this ability or to lack it altogether.

Where fictional-x_i violates $\mathrm{IWV}_{\mathrm{revised}}$, the charge would be that the fiction (presumably, owing to the fiction's creator) is, at best, insensitive to the moral ambiguity surrounding the meaning of the fiction and subsequently insensitive to, or at least willing to disregard, the likelihood that the fiction will be taken by some (many), based on a reasonable interpretation of the fiction, to be endorsing an immoral worldview, even if this is not the intention of the fiction and therefore its creator. As a consequence, the fiction's creator is being insensitive to the *feelings* of those who, if confronted by the fiction, would likely be troubled, even distressed, owing to the ambiguity of the message, irrespective of any distress caused by the depiction itself (i.e., in terms of graphic realism, intensity, and duration). The moral insensitivity I am suggesting, here, is expressed more bluntly by Patridge, who claims, in the absence of mitigating circumstances (to be discussed), minimally, it is reasonable to believe that what such fiction constitutes, owing to its ambiguity, is "a thumbing of one's nose at a requirement of solidarity with the victims of oppression" (Patridge, 2011, p. 310). That said, I hesitate to ascribe to the fiction an immoral status. Instead, by incorporating the notion of moral insensitivity into $\mathrm{IWV}_{\mathrm{revised}}$, rather than immorality, we get:

> $\mathrm{WV}_{\mathrm{insensitive}}$: Where a reasonable interpretation of a fiction, taken as a whole, cannot establish whether the fiction is merely depicting rather than endorsing an immoral worldview, then any depiction of immorality from within the narrative congruent with that worldview is, minimally, morally insensitive.

The vulnerability of fiction to a charge of moral insensitivity increases in proportion to its seeming insensitivity to depictions that appear to convey incorrigible social meaning (e.g., the depiction of an African American navigating a field of watermelons), but, minimally, it is still applicable to all forms of immorality. A charge of moral insensitivity simply has less force in cases where the fiction depicts, ambiguously, at least in terms of associated endorsement, organized crime, say, or confidence tricksters, owing to their lack of incorrigible social meaning.[9] In other words, the seeming endorsement of "gangster" culture is vulnerable to a charge of moral insensitivity given the immoral activities associated with "gangsters" and therefore given the number of victims of their immoral behavior, even if "gangster" behavior has never been previously valued, and even if such seeming endorsement is not as insensitive, owing to a lack of incorrigible social meaning as, say, rehashing the *Black and White Minstrel Show*, in which white performers "black up" to caricature persons of African descent.

In accordance with IWV and $\text{WV}_{\text{insensitive}}$, where fictional-$x_i$ depicts violence toward women (for example) or members of a minority group (e.g., homosexuals), or any other form of immorality, the following applies:

(i) Where fictional-x_i occurs within the context of a narrative that endorses immorality and is congruent with that narrative (whether it is the view that, inter alia, violence toward women or homosexuals is acceptable or that organized crime or being a "con artist" is likewise acceptable), given IWV, one would be justified in concluding that fictional-x_i is immoral.
(ii) Where ambiguity exists over whether fictional-x_i satisfies IWV, in accordance with $\text{WV}_{\text{insensitive}}$, and in the absence of mitigating circumstances, one is nevertheless justified in declaring that fictional-x_i is, minimally, morally insensitive.

Satisfying (ii) would make the fiction morally troubling; satisfying (i) makes it morally dangerous (Mullin, 2004). I will return to the issue of what makes $\text{WV}_{\text{insensitive}}$ morally troubling in chapter 9 when discussing suberogatory actions. In the meantime, where one does not wish to fall foul of $\text{WV}_{\text{insensitive}}$, and, therefore, where one cares about not being accused of moral insensitivity, what other motive might there be for depicting immorality in a manner that makes the fiction vulnerable to $\text{WV}_{\text{insensitive}}$? It could be that one is prepared to risk being misconstrued as endorsing immorality if the benefit one seeks is edification (i.e., engaging with *this* depiction in *this* way is a good or perhaps the best way to learn from the fiction). One may believe that using fiction to create a shocking or disturbing experience, perhaps even as a work of art, is a way to edify, morally, and/or cultivate aesthetic appreciation, as well as a way to entertain (Devereaux, 2004). Let us therefore examine the role of moral edification and the cultivation of aesthetic appreciation as a way to mitigate (potential) moral insensitivity, starting with the former.

MITIGATION IN THE FORM OF MORAL EDIFICATION

When contemplating the meaning of a narrative featuring depictions of fictional immorality, it is unrealistic to demand that the possibility of misunderstanding is zero, or as close to it as to make it negligible. Nevertheless, in an attempt to reduce misunderstanding—say, by engaging with depictions of fictional immorality as the creator intended, rather than in a manner indicative of one's misinterpretation—one could employ the technique of *hermeneutic recalibration* (Liao, 2013; see also Liao & Gendler, 2015). On those occasions when one is initially puzzled by the fiction, perhaps by design (i.e., if it is the creator's intention to unsettle the audience and/or force them to consider a different perspective or perhaps a number of different perspectives, including, say, moral inversion), through hermeneutic recalibration, one eventually arrives at *a* relatively stable interpretation, grounded on the understanding and acceptance that what is depicted, as ambiguous as it may have at first appeared, is a *fiction*, possibly with its own "in-house" rules, including (perhaps) a nonrealistic fictional component or feature (e.g., the moral inversion found in *Catch-22*).[10] Where it is not the intention of the fiction and therefore its creator to endorse immorality, the likelihood of misconstruing fictional-x_i rests on the extent to which one interprets the nonrealistic components—say, in the case of moral inversion—as persuading the audience to believe rather than make-believe (through their fictive imagining) the moral inversion depicted. Or, where the depiction aligns more with realistic fiction, the underlying moral message of the confronting fictional immorality needs to be one's focus rather than the make-believe one is invited (again, through one's fictive imagination) to engage with, however horrific and unsettling that may be. As Six, the director of *The Human Centipede 2*, remarked, after criticism over depictions within his film: "Shouldn't a good horror film be horrific."[11]

Where there is a reasonable likelihood of an audience misconstruing the meaning of the fiction, irrespective of whether hermeneutic recalibration would ease this misunderstanding, perhaps any negative judgment about the wrongness of the depiction and, subsequently, the intent of the fiction's creator—to depict fictional immorality regardless of the risk of the audience mistaking its meaning—can be mitigated by the message the creator of the fiction is trying to convey and therefore by what he or she might hope the audience will *learn* from the fiction taken as a whole, or come to appreciate, aesthetically (although I will save discussion on this latter possibility until the next section). Perhaps one way the risk can be off-set is by moral edification, and perhaps it is the opinion of the creator of the fiction that this is *a* way—maybe even the best way—to edify the audience (see Carroll, 2002; Currie, 1995; Depaul, 1988; Kieran, 1996).[12] To borrow from Mills (2000), albeit in a slightly different context than originally intended: "It is also relevant whether the writer *intends* harm [*qua* misunderstanding] or whether any harm [*qua* misunderstanding] she causes is merely the foreseen but regretted by-product of other morally legitimate activity" (p. 197; emphasis in original).

However unpleasant one's experience may be (see the brief discussion on the paradox of negative fiction in chapter 8), there is value in what we can learn about ourselves and/or society by engaging with a fiction like Bret Easton Ellis' *American Psycho*. The novel (and later the film) depicts the actions of a fictitious serial killer in brutal detail but, importantly, from the point of view of someone whose emotional engagement with his victims and concern over the viciousness of his actions are as superficial as his interest in the latest fashion accessory or design of business card. For Oxenbøll (2017), depictions of fictional violence and other taboos create learning spaces for the audience, including opportunities for engagement *at a safe distance*. They provide an audience with the opportunity to discern meaning from the mayhem depicted or to reflect on how violence constitutes a perversion of normal social structures and cohesion. Fictional immorality of this kind satisfies Mills' (2000) claim about fiction more generally: that much of it is more than trivial entertainment, and the best of it can change lives. Gaut (2007) likewise comments on the epistemic value of art (much of which satisfies my definition of pure fiction) in terms of what it can teach us about morality: "In assessing the ethical value of art we are assessing the ethical quality of the *point of view*, cognitive and affective, that it takes toward certain situation" (p. 9; emphasis added). Oxenbøll adds, however, that tolerable depictions of violence make for easier learning, thereby recognizing that depictions of violence and other taboos that are too lifelike may cause a great deal of discomfort to the audience. Such discomfort may result in disengagement with the fiction. However, it may have been created with just this form of confronting education in mind. A view echoed by Miller (1975) when commenting on the violence in Sam Peckinpah's *The Wild Bunch* and *Straw Dogs*:

> [The violence in these films] need not prove that Peckinpah wants a lot of people to jump up and down in bloodthirsty glee and then go home and kick the dog; nor does it mean that he wants them to work out their frustrations in the theatre and go home mellowed . . . Peckinpah wants to worry his audience—if they hoot happily at suffering, that's their problem. (p. 3)

To illustrate the idea of confronting but educational fiction further, at the time of its release, the brutal rape scene depicted in Gaspar Noé film *Irreversible* was the focus of much debate among reviewers. Mitchell (2003, p. 1), for example, states: "It's no doubt that he [Gaspar Noé] wanted to make a film that navigates the fine line between noxious and obnoxious," and in doing so, he has presented us with one of the "most gruesomely detailed rapes" ever to be witnessed on the screen. The camera takes in the scene "without looking away—that's left up to you," he tells the reader (*ibid.*). In a similar vein, McAllister (2006, p. 1) asks whether the violence depicted in *Irreversible* was necessary or gratuitous. In response, he states: "I would argue that these scenes are so horrifically convincing that they become virtually impossible to watch, and this is the effect that violence should have." In a similar vein, Katie Mitchell, theater director of the Sarah Kane play *Cleansed*, which features enactments of graphic violence, admonishes British audiences for being afraid of exploring the violence of atrocity, while declaring that we need to get better at confronting difficult issues.[13]

Fictional portrayals of the kind described are not, however, without risk, but they can be morally valuable precisely because they help to dispel the notion that evil is somehow outside of us (Koppelman, 2005). The "risk" is either (and more likely) the risk of misconstruing the meaning of the fiction—that it is endorsing sexual violence toward women (for example), with the moral condemnation and disengagement that follow, rather than educating us to the reality and horror of it—or, relatedly (although less likely, I feel), the increased risk of individual and societal harm occurring in the form of antisocial behavior, attitude, or affect either as a consequence of the misconstrued message or simply through engaging with this type of fictional immorality (an issue I will discuss in more detail in chapters 5 and 6 in the context of harm and altering moral outcomes). Or it could be, as Oxenbøll (2017) cautions against, that the depiction is too realistic, resulting in a turning away from the fiction and subsequent failure to embrace the learning space and hence intended message. As a critic of the aforementioned play, *Cleansed*, asks: Do we really need to witness graphic depictions of violence in order to reflect upon the darker side of human nature?[14]

In concluding this section, if, on account of a reasonable interpretation of the fiction, one judges that the meaning of the fiction is unclear and could be misconstrued, then, even if the fiction was not created with the intention of endorsing immorality, and so cannot be said, by design, to be complicit in endorsing that which is immoral, the fiction (and its creator) could nevertheless be accused, with some justification, of being minimally complicit in allowing misunderstanding, and therefore of being moral insensitive to this possibility and the distress it may cause. More than this, one could be accused of the *wilful fostering* of misunderstanding, along with the subsequent charge of indirectly bolstering an immoral worldview—such as "violence toward women and minorities is okay"—as a consequence of this willful ambiguity. However, it may be that one is prepared to risk audience misinterpretation, and even accusations of impropriety of the kind presented above, if one believes that the risk and hence the potential cost of any misunderstanding are outweighed by the benefits to the audience of engaging with *this* fiction *in this way*. A factor that might mitigate the risk of being accused of violating IWV, or even the lesser indiscretion of failing to meet the requirements of $\mathrm{WV}_{\mathrm{insensitive}}$, is the intent to use the fictional immorality as a means of morally educating one's audience, in this case through the use of potentially harrowing and morally provocative depictions.

THE AESTHETICIZATION OF VIOLENCE AND OTHER IMMORAL ACTION

Consider the following commentary by Lang (2019) on the violence depicted in films by Quentin Tarantino.

> Violence in Tarantino's films is virtually never reflective of what violence looks like in real life: It is unnatural, unrealistically bloody, and heavily stylized. Severed bodies magically contain more than the standard 5.5 liters of blood in order to supply dramatic splatters; limbs regularly go flying; and a buddy accidently shot in the face is an inconvenience to the shooter, rather than a horror . . . Tarantino has likened his slasher scenes to doing what dance sequences do in musicals: function as choreographed spectacles that are intended to be relished. (p. 1)

Depiction of fictional violence as *choreographed spectacles* that are *intended to be relished*, as Lang puts it, is as succinct a way as any to capture what is meant by the *aestheticization of violence* (see, for example, Bacon, 2015; Prince, 1998; Schneider, 2001; Sheehan, 2013). In accord with this view, perhaps the purpose of the fictional violence depicted (or other immoral acts) is not to educate the audience about the true nature of what they are witnessing—say, by compelling them to adopt or at least contemplate a particular moral stance—but to afford a means of aesthetically appreciating the way the violence has been *depicted*, which may lack a degree of realism (although the aim may be to achieve both moral education and aesthetic appreciation; see Cahn's (1974) discussion on Kubrick's *A Clockwork Orange*). Rather than being led to understand what violence really is, through highly stylized depictions of immorality, the audience is instead invited to appreciation what they would likely find abhorrent and reject if witnessed for real. As Symonds (2008) notes in relation to the kind of martial arts violence Tarantino pays homage to in his *Kill Bill* films:

> Violence can be the medium through which fantasy about the choreographed fighting body can thrill, amaze and uplift the viewer by creating a sense of the

> body liberated from its physicality. . . . [Thus, fight] scenes transform violence into a special effect dreamscape based on inspirational images of the body in f(l)ight [flight and fight]—images that uniquely depend on depriving violence of its visceral authenticity to create a sense of audience wonder. (pp. 151–152)

In short:

> [One should] delight in the grace of martial arts violence rather than be distressed by it. (*ibid.*, p. 176)

I am thus reminded of the "Mexican standoff" that takes place in Sergio Leone's *The Good, the Bad, and the Ugly*. The cinematography, alongside Ennio Morricone's music and what are effectively carefully choreographed dance steps by the three protagonists, positioning themselves for the fight to come, combine to create a prelude to a depiction of an immoral act I always relish. Similarly, and again, in accordance with Symonds (2008), the slow motion, gravity-defying fight scenes in *The Matrix* give the violence an embodied grace that appears surreal.

In support of the aesthetic quality of certain depictions of fictional violence, Kreider (2008) argues that "most people would agree that an artistic context does justify or at least mitigate some things that we would otherwise find objectionable" (p. 153). A view shared by Poole (1982) when stating that the depiction of "morally and emotionally shocking situations might be tolerable if the author *intended* to create a work of art" (p. 40; emphasis added). Indeed, as Symonds (2008) notes when discussing the impact of certain depictions of violence that do shock (e.g., the scourging of Jesus of Nazareth in Mel Gibson's *The Passion of the Christ*: "Creating blood effects on stage and screen is always about 'effects' rather than authenticity" (p. 169), adding that the power of the "effect" to disturb lies with the depictions copiously detailed and virtually enhanced intensity, rather than its realistic authenticity.

Kreider, Poole, and Symonds' comments support the importance of "artistic license" and therefore the merits of the aestheticization of violence as a mitigating factor able to protect fictional-x_i from a charge of violating $WV_{insensitive}$. Not only that, but their comments are not incompatible with the idea of fictional immorality as moral educator; although such a view must be weighed against the fact that in order to educate (in a nontrivial sense) through fictive imagining, one must not deviate too far from reality, at least in terms of the fixed norms one is seeking to endorse or challenge (Gaut, 2007). As the audience, then, we are left to consider the nature of the fictional world we have been invited to enter and the manner of that invitation (Dillon, 1982). Good literature and art (including films and other media) typically invite us to consider the point of view proffered by its creator and will perhaps even challenge our own views (Young, 2000); it does this irrespective of whether we take up the challenge or succeed in being edified. Of course, it may be that we simply enjoy the aesthetic quality of what is depicted and the manner of its depiction, and this is reason enough for our engagement with the fiction.

In chapter 7, I discuss the morality of enjoying fictional immorality—in terms of its entertainment value—and whether the nature of one's enjoyment mitigates or contributes to fictional-x_i's vulnerability to the charge of violating $\text{WV}_{\text{insensitive}}$.

CONCLUSION

In accordance with IWV, I have argued that a fiction that endorses immorality (*qua* an immoral worldview) is immoral because, in effect, one is using the fiction to promote the (immoral) beliefs one is endorsing (e.g., some races of people are inferior to others). I have also argued that while upholding IWV is a worthy pursuit, it is likely to be difficult to apply owing to ambiguity within the fiction. As a consequence, I entertained the possibility of "erring on the side of caution," by adopting a "reasonable interpretation" approach to judgments about fictional immorality ($\text{IWV}_{\text{revised}}$): something I ultimately found to be unsatisfactory, although I did favor an argument for moral insensitivity ($\text{WV}_{\text{insensitive}}$). I also discussed how the risk of misunderstanding the meaning of fictional content should be weighed against the potential educational and aesthetic benefits of engaging with the fiction, and how moral edification and aesthetic appreciation might act as mitigating factors when examining the case for immorality.

In finishing this chapter, it might be that, irrespective of whether a fiction endorses an immoral worldview, and therefore irrespective of the immorality of the meaning of the message conveyed by the fiction, or whether one misconstrues or not the meaning of the fiction and/or its aesthetic pursuit, engaging with depictions of fictional immorality causes harm to oneself or others, or at least increases the risk of harm occurring. This issue will be taken up in the next two chapters.

NOTES

1. https://www.adl.org/sites/default/files/documents/assets/pdf/combating-hate/Racist-groups-use-computer-gaming.pdf (accessed July 29, 2019).

2. Minorities are portrayed as gangs of "subhumans," often toting machine guns.

3. It is unremarkable to point out that, historically, moral standards have varied, even in recent years. In the 1955 film *The Seven Year Itch*, for example, which features a married man who fantasizes about an affair with his neighbour (played by Marilyn

Monroe), marital infidelity does not occur because production codes at the time prohibited the writer and director, Billy Wilder, from filming a comedy that featured adultery.

4. One could also present the premise as follows: S *believes* that fictional-x_i is congruent with the immoral worldview that (in accordance with [b]) S believes is being endorsed by the fictional narrative taken as a whole. This would be an epistemic rather than ontological claim. This difference does not change the argument I presented, here, however.

5. A different way of presenting $\text{IWV}_{\text{revised}}$ is: where a *reasonable* interpretation of a fiction, taken as a whole, does not preclude the possibility that the fiction is endorsing an immoral worldview, then any depiction of immorality from within the narrative, congruent with that worldview, is immoral. Precluding the possibility may be too severe a requirement, however.

6. To be fair to $\text{IWV}_{\text{revised}}$, it is not attempting to be infallible.

7. In the United States in 2003, Los Angeles officials asked manufacturers, suppliers, and contractors of computer hardware to refrain from using the terms "master" and "slave" to refer to types of equipment, adding that such terms were unacceptable and offensive (CNN, 2003).

8. In a similar vein, "slasher" horror films (e.g., *Halloween*, *Friday the 13th*) have been criticized for depicting a disproportionate amount of violence toward women. Sapolsky et al. (2003), however, analyzed the content of 1980s and 1990s slasher films and found this not to be the case. In fact, males are targeted more often, although there were more prolonged exposure of females exhibiting fear and distress.

9. Owing to the contingent nature of incorrigible social meaning, ignorance can sometimes be a defense. If one is simply not aware of a particular association that has incorrigible social meaning for a given society (or subgroup) and, importantly, it is unrealistic to expect that one should have been aware of this, then one is hardly expressing moral insensitivity, even when one's action may have caused distress and/or offence. Once the incorrigible social meaning is revealed, however, if one chooses to ignore this, then the previous mitigation would be dispensed with.

10. Liao (2013) notes that hermeneutic recalibration is often required when reading magical realist novels, but there is no reason for the technique to be restricted to this style of fiction.

11. https://www.telegraph.co.uk/culture/film/film-news/8563325/Human-Centipede-II-director-says-banned-film-is-art.html (accessed May 28, 2019).

12. Booth (1964) makes a similar point about moral edification in relation to stories found in the Bible. "When we read the many . . . specific accounts of sexual abuses that the Bible contains—of seduction, incest, sodomy, rape, and what not—we do not put the Bible on the list of banned books, because we know that the context requires an honest treatment of man's vices" (p. 158). I use this example for illustrative purposes, only. Its inclusion is not meant to suggest that the Bible contains, in whole or in part, depictions of fictional immorality. Equally, I am not claiming that it does not.

13. https://theconversation.com/how-far-should-we-go-when-depicting-violence-55560 (accessed July 19, 2019).

14. Again, see https://theconversation.com/how-far-should-we-go-when-depicting-violence-55560 (accessed July 19, 2019).

Chapter 5

The Harm of Fictional Immorality

INTRODUCTION

An accusation often leveled against depictions of fictional immorality, especially those involving violence, is that engaging with them increases the likelihood of *harm* occurring, either to oneself or others, or both.[1] Should this accusation be justified (i.e., should it achieve empirical-based *consensus*; see chapter 6 for further discussion on this point), then, for this reason alone, our engagement with fictional-x_i warrants moral appraisal. This is because causing or being caused to suffer harm has moral significance for us (Hanser, 2008, 2019). We are typically motivated to prevent or avoid harm or to assist those who have been harmed; but more than this, we accept that this is how we *ought* to behave—in an obligatory sense—when confronted with harm or its possibility. With regard to fictional immorality, then, the following "immorality condition" (IC) reflects this normative approach:

IC_1 Engaging with fictional-x_i is immoral *if* it causes harm.

The aim of this chapter is to present a qualified defense of IC_1. I begin by discussing *why* causing harm is immoral. I then consider what it means for someone to suffer harm or for someone or something to cause harm. To do this, I examine different accounts of what must be satisfied in order for *harm* to occur. After that, I stipulate the conditions under which fictional-x_i can be said to cause harm, either directly or indirectly. Finally, while speculating over the absence of a direct or indirect causal connection between engaging with fictional-x_i and harm, I assess the value of adopting a precautionary approach to harm based on the risk—*qua* increased likelihood—of harm occurring. What will not feature in this chapter, however, is discussion on the merits of IC_1 based on the findings of research that has looked at the relationship between fictional violence and harm. Answers to questions such as “What *evidence* is there for a causal connection between engaging with depictions of fictional immorality and harm?” or “What *evidence* is there for the *increased likelihood* of harm occurring after engaging with depictions of fictional immorality?” will be postponed until the next chapter.

Before continuing, however, it is worth noting that depictions of fictional immorality appear in many forms: both in terms of the immorality depicted (e.g., murder or torture or racism) and the medium in which these depictions appear (e.g., a film compared to a cartoon, or literature compared to a song lyric, and so on). To say that the various connotations of fictional-x_i available across all media form a homogenous group is to concede only that they do so, nominally, in virtue of the fact that they are all depictions of *fictional immorality*. Such nominal homogeneity does not necessitate similarity in the amount of harm associated with the act of engaging with these depictions, which may vary depending on the specific immorality depicted or the medium in which it occurs. Whether this is the case will be determined by gauging the strength of empirical evidence, rather than a priori reasoning; but, to reiterate, discussion on this matter is postponed until the next chapter.

THE IMMORALITY OF HARM

At the start of the chapter I noted how, typically, we are motivated to prevent or avoid harm, or to assist those who have been harmed, and how this is something we feel obliged to do because causing or being caused to suffer harm has moral significance for us. Importantly, though, the proposition “harm is morally wrong” does not contain an analytic truth (Bradley, 2012). The moral wrongness of harm is not revealed through an analysis of the term “harm.” Instead, research supports the view that we develop moral associations based on our natural aversion to harm in the form of a disgust response (see Young & Whitty, 2012, Chapter 3 for a more detailed discussion), but also that we have an aversion to harm-inducing events, even when these are simulated rather than real (e.g., stabbing someone with rubber knife; see Cushman et al., 2012). Where harm is perceived to have occurred, we often associate wrongdoing and, subsequently, immorality with the event (I am excluding natural events—inter alia, earthquakes, floods, wild animal attacks—from this claim).

In an attempt to refine our understanding of the aversion we exhibit toward harm-inducing events, Miller and Cushman (2013) distinguish between outcome-based aversions, which concern the consequence of an event (where this is perceived to be harmful), and action-based aversion, which is to have an aversion to the act itself, or even to the idea of it. The example they give in the latter case is punching one's mother in the face. We have a strong aversion to this act because of its harmful consequence but also because of what is involved in the act itself (i.e., clenching one's fist, raising it up, and directing it forcefully toward the face of one's mother). Likewise, Schein and Gray (2017) report that when presented with the sentence "The man intentionally *gished* the little girl, who cried," even though "gished" is not a real word (in English), we typically judge the man to have done something wrong. Causing a vulnerable person—a "little girl"—to cry (presumably tears of sorrow rather than joy) is to cause the girl to suffer harm, and doing this intentionally leaves the man vulnerable to the charge of acting immorally.

According to Gert (2011), aversion to harm is a sign of our rationality: for all *rational* people are motivated to avoid harm unless they have a reason not to. One's reason for not avoiding harm may itself be rational, of course—for example, self-sacrifice in defense of one's family—or irrational, owing, say, to some mental disturbance; but importantly, for Gert, if rational people are motivated to avoid harm and all people, to a greater or lesser degree, are vulnerable to harm, then it seems rational to develop a system designed to protect us from harm. This system, Gert informs us, produces our common morality (see Gert, 2004). In a similar vein, Mill has the following to say: "The moral rules which forbid mankind to hurt one another . . . are more vital to human well-being than any maxims" (Mill, 1957 [1861], p. 68). In fact, for Mill (2005 [1859]), the only justification for restricting one's liberty is if one's liberty causes harm to another, at least where the harm constitutes an overall harm, rather than a *pro tanto* harm. Thus, cutting someone's flesh with a knife is harmful to the person cut, but if performed by a surgeon (for example), as part of an operation to improve the patient's condition (i.e., save their life or otherwise improve their health), then the act of cutting flesh, while always harmful to some extent, is nevertheless not absolutely harmful.[2] Consequently, it should be classified as a *pro tanto* harm, especially as it is not the intention of the surgeon to make the patient worse off. Where the latter outcome transpires, however—that is, where the patient is made worse off—whether intentionally or not, then the patient would have suffered harm that is intrinsically bad for them and, for this reason, suffered a harm that is immoral (Bradley, 2012; Harrosh, 2012).

WHAT IS IT TO HARM OR BE HARMED?

We have seen how harm, at least in the sense in which we are interested (i.e., not a *pro tanto* harm), has a negative effect on the person harmed, insofar as it makes things worse for them.[3] According to Feit (2015), a person is harmed *if* and *only if* they are made worse off (recall that, by being worse off, S is in a state that is intrinsically bad for him; see Bradley, 2012; Harrosh, 2012).[4] By contrast, however, I intend to argue that while "being worse off" (relative to some other condition or time) is sufficient for harm, it is not necessary. To understand why, it is important to note that approaches to harm differ with regard to how this negative state (this intrinsic badness) is measured. Comparison theories hold that the harm incurred—the negative effect—is relative to either S's prior state (temporal comparison) or what S's state would have been had the harm-inducing event not occurred (counterfactual comparison). Noncomparison accounts, in contrast, maintain that S is harmed whenever she breaches a threshold beyond which one is said, in a categorical sense, to have been harmed.

Comparison Approaches

The temporal-comparison approach has intuitive appeal. It holds that one suffers harm if one is worse off at t_2 compared to t_1. Thus, where S is assaulted on his way back from a meeting with a client, he is worse off on account of the assault (at t_2) than he was before the assault (at t_1). There are problems with this approach, however, as the following example illustrates. By way of a thank-you from a satisfied client, S receives a gift card containing a lottery ticket (the client is closing their business and so will never see S again). Unbeknownst to all, the lottery ticket is a winning ticket. Now, rather than posting the gift card or delivering it herself, the client asks S's colleague to pass it on. The colleague agrees. He is, however, jealous of S's business success and, because of this, throws the gift card away. S is forever unaware of his former client's kind gesture, and they are forever unaware that he did not receive the gift card. Despite the action of his colleague, at t_2 (after the gift card has been thrown away), S is no worse off than he was at t_1: he did not have a winning lottery ticket then and he does not have one now. According to the temporal-comparison approach, S has not been harmed by his colleague's action. This outcome is counterintuitive.

The counterfactual-comparison approach (Feinberg, 1984), in contrast, would conclude that S was harmed by his colleague's action (let us call this event E). The approach holds that S is harmed by E if he is worse off on account of E than he would have been had E not occurred (see Carlson, 2019, for a recent critique). Had E not occurred, not only would S have been aware of his former clients' kind gesture, he would also have been in possession of a winning lottery ticket.

Again, this approach has intuitive appeal. Moreover, it avoids the problem faced by the temporal-comparison approach. Objections have been raised, however. To illustrate: Where S has had his wallet stolen, the counterfactual-comparison approach would claim that S is harmed because he is worse off than he would have been had his wallet not been stolen. This seems perfectly reasonable. Suppose, however, that on realizing that his wallet has been stolen, S suddenly halts. The fact that he stops enables him to avoid a ricocheted bullet (from an armed robbery across the street) that would otherwise have injured him. If his wallet had not been stolen, he would not have halted (at least not for this reason and unlikely for any other) and, as a consequence, would have been hit by the bullet. The example challenges the counterfactual-comparison approach because it shows that S was not in fact worse off by having his wallet stolen than he would have been in the counterfactual condition (in which his wallet is not stolen) because he would have continued walking and been hit by a ricocheted bullet. If S is not worse off in the actual condition, compared to the counterfactual condition, then he has not been harmed by having his wallet stolen. Again, this outcome seems counterintuitive.

We could, of course, respond with an even more convoluted scenario. When S stops abruptly after realizing that his wallet has been stolen, although, as a consequence, he fails to be hit by the ricocheted bullet, which would have wounded him but not fatally, he is instead hit by a different bullet (from the same armed robbery) that kills him. He is thus worse off than he would have been in the counterfactual condition (where his wallet is not stolen and he is only wounded, not killed). In this scenario, S *is* harmed by having his wallet stolen, rather than not having it stolen, but only because he is also shot and killed rather than shot and wounded!

To avoid the need for overly convoluted scenarios, Feinberg (1984) argues that the *counterfactual* harm must be *reasonably foreseeable*. Thus, according to Feinberg (see also Purshouse, 2016, for a detailed discussion), it is illegitimate to declare that S is not harmed by having his wallet stolen owing to the fact that, in the counterfactual condition, he is injured by a ricocheted bullet. This move is illegitimate because the counterfactual injury is not reasonably foreseeable. But even if we accept Feinberg's requirement, what is irksome about each of the "wallet" scenarios—regardless of whether they involve S having his wallet stolen, or (let us allow for the sake of the next point) not having it stolen but being shot and injured, or having it stolen and being shot and killed—is that, in each case, S would be worse off than he was *prior* to any of these possible happenings. Irrespective of which of the three scenarios befalls S, he would be harmed at t_2 because things would be worse for him at that time than they were at t_1. This realization brings us back to the temporal-comparison approach; but, as noted, this approach has problems of its own.

A possible way forward, if one does not wish to commit oneself solely to the counterfactual-comparison approach, is to make temporal and counterfactual comparisons individually sufficient but not necessary. Thus:

A S is harmed *if*, on account of E, things are worse for S at t_2 than t_1.
B S is harmed *if*, on account of E, things are worse for S than if E had not occurred.

In the discarded gift card example, A is not satisfied; but, as this is a sufficient condition only (not a necessary one), some other conditions may suffice. In this case, B is satisfied. As previously noted, S would have been better off if his colleague had not thrown the gift card away. Conversely, in the example in which S's wallet is stolen, and it is contrived that B is not satisfied (owing to the fact that S would have been shot and injured if his wallet had not been stolen), A is satisfied. In fact, recall in the wallet example, at t_2, S is worse off in *all* of the scenarios presented, compared to t_1. As individually sufficient conditions, but not necessary ones, each comparison approach is able to explain what could qualify as making S worse off on a given occasion. Moreover, in the simple case of S *just* having his wallet stolen, both A and B are satisfied: S is worse off than he was before and worse off than he would have been if his wallet had not been stolen. The fact that both apply, on this occasion, strikes an intuitive chord. Therefore, each comparison approach tells us something about what would constitute being *harmed.*

The Noncomparison Approach

Hanser (2008) acknowledges that to suffer harm is to be worse off (than not suffering harm), and therefore to cause harm is to cause someone to be worse off, but rejects temporal or counterfactual comparisons in favor of a noncomparison approach. Using the example of blindness, he argues that S is harmed by blindness even if S suffers from congenital blindness and there was never a prior time when he was able to see or a counterfactual world in which he does not suffer from congenital blindness. Echoing Shiffrin's (2012) view, Hanser (2008) holds that, in being blind, S is harmed in an absolute sense rather than relative to an earlier time or counterfactual condition.[5] Put differently, for Hanser (2011), S is harmed *if* and *only if* he is in a state that is simply bad for him, and therefore where the occurrence of harm is not determined by the condition of associate states.

Thomson (2011), however, objects to Hanser's (2008) claim that "being blind" is a noncomparative harm. She argues that the state of "being blind" is not noncomparatively bad, and therefore harmful, if one is a bat, for example (although it has to be said that bats are not completely blind). The same could be said of the star-nosed mole or, perhaps (as a better fit for a totally blind creature), the blind huntsman spider (*Sinopoda scurion*). Rather, being blind can be thought of as harmful relative to X *if* X's nature is not to be blind. It is thus harmful in comparison to the normative biological state of the species to which the organism belongs. For something to be noncomparatively harmful, Thomson tells us, not only must the organism fare badly irrespective of a prior or counterfactual condition/state of being, rather than relative to it (just as Hanser attests), more than this, it must not fare badly relative to its normative biological state.[6]

Thus, where S has been confined to a two-square-meter cell for the whole of his life, and has therefore been deprived of his liberty from birth, he has been and continues to be harmed. This harm is not relative to his normative biological state, which has nothing to say about liberty; rather, it is absolute harm. It is absolute harm because the harm remains even when temporal and counterfactual comparisons fail to yield harm.[7] To illustrate: when S is eventually "upgraded" to a ten-square-meter cell, he is arguably better off than he was before, or at least not worse off, as a result of the change at t_2 compared to t_1. Given this, A (introduced earlier; see again below) does not apply:

A S is harmed *if*, on account of E [the "upgrade"], things are worse for S at t_2 [after the "upgrade"] than t_1 [before the "upgrade"]

According to A, S is not harmed by his switch to a larger confinement cell because he is not worse off at t_2 than t_1 (relatively speaking), although, it has to be said that he still fares badly, overall. Equally, because the occurrence of the "upgrade" does not make things worse for S than if it had not occurred, he is not harmed by his change of circumstance. Had the upgrade not occurred, S would have remained confined to a smaller cell. Given this, B (again, introduced earlier) does not apply either:

B S is harmed *if*, on account of E [the "upgrade"], things are worse for S [with the "upgrade"] than if E [the "upgrade"] had not occurred.

Because S is *not* worse off with the "upgrade" (relatively speaking) than if the "upgrade" had not occurred, as far as B is concerned, S is not harmed. Yet, do we really want to say that being confined to a ten-square-meter cell is *not* harmful because one is not worse off relative to a counterfactual condition not occurring, or relative to a prior time in S's history?

Combining Approaches

A and B are considered sufficient for harm, but not necessary. In the confinement scenario, neither A nor B is satisfied, and yet S *is* suffering harm. A further—noncomparison—condition must therefore be added that, like A and B, is sufficient but not necessary for harm.

C S is harmed *if*, on account of E, S fares badly, irrespective of temporal, counterfactual, or normative biological comparisons.

The fact that the noncomparison condition is sufficient for harm means that S can be harmed even when A and B fail to be satisfied, and therefore indicate no harm is occurring (relatively speaking). The noncomparison condition is not necessary for harm, however, because relative harm is still harm, and so S can still be harmed, relatively speaking (by satisfying A or B), even if S is not harmed according to C; that is, if S fares badly only with respect to temporal, counterfactual, or normative biological comparisons and not in some absolute sense or satisfies C in addition to and therefore irrespective of A or B.

With respect to the former possibility for harm, where S is temporarily deprived of liberty by a nefarious colleague—by being locked in a room against his wishes for a few hours—then either A or B is satisfied in the absence of satisfying C. By contrast, and with respect to the latter possibility for harm, in the lifetime confinement scenario discussed earlier, S is already suffering harm in virtue of breaching a threshold beyond which one is being harmed—for example, having one's liberty denied or restricted (a harm, to reiterate, that is not relative to biological normativity and is in keeping with Gert's, 2004, basic harms)—and this remains the case even when S is transferred to a different size cell because, irrespective of the size of the cell, S is still confined and so still fares badly. In other words, where E constitutes a lifetime confinement to a cell, S fares badly. Alternatively, where E constitutes an "upgrade" to a larger cell, or downsizing to a smaller one, either way, S remains confined and so still fares badly. That said, even where this is the case, S could still find himself worse off *relative to* the following temporal or counterfactual conditions:

- (Temporal comparison) Where S is harmed because all of his life he has been confined to a ten-square-meter cell, he is still worse off (suffers greater harm) if he is suddenly confined to a two-square-meter cell.
- (Counterfactual comparison) Where S is harmed because all of his life he has been confined to a ten-square-meter cell, he is still worse off (suffers greater harm) on account of being transferred to a two-square-meter cell than if the transfer had not occurred.

THE CONDITIONS FOR HARM

Despite being presented as contrasting approaches in the literature (Hanser, 2008, 2019; Rabenberg, 2015; Shiffrin, 2012), I nevertheless take comparison and noncomparison approaches to be compatible, if treated as sufficient conditions rather than necessary ones. Conditions A–C each provide a description of a state of affairs that, if any is satisfied, is sufficient for someone to have been harmed. When presented in this way, it is possible for the magnitude of harm to increase or decrease relative to a prior time in the subject's history or a counterfactual condition, even while maintaining an overall harm status (*qua* faring badly). When combined, the following "harm condition" (HC) is produced with regard to the suffering of harm:

> $HC_{suffering}$ S is harmed *if*, on account of E, S fares badly irrespective of temporal, counterfactual, or normative biological comparisons, or *if* things are worse for S at t_2 compared to t_1 or than *if* E had not occurred.

Applying $HC_{suffering}$ to the act of engaging with depictions of fictional immorality, we get:

- S is harmed *if*, on account of engaging with fictional-x_i, S fares badly irrespective of temporal, counterfactual, or normative biological comparisons, or things are worse for S after engaging with fictional-x_i compared to before the engagement or than *if* engaging with fictional-x_i had not occurred.

How might S come to suffer harm in the manner described above: that is, be worse off based on a temporal or counterfactual comparison, or fare badly in a noncomparative way? Recall that, for Mill (2005 [1859]), the *only* justification for restricting a person's liberty is if, as a consequence of exercising their liberty, they harm another (recall, also, that we are not talking about *pro tanto* harm, here). Typically, restricting someone's liberty makes things worse for them (compared to before their liberty was restricted or compared to a counterfactual condition in which it is not restricted), and, irrespective of comparisons, it is an imposition that causes one to fare badly. So, if making something worse for another person—that is, harming them—is sufficient justification for making things worse for the one who did the harming (i.e., restricting their liberty), then one way in which engaging with fictional-x_i could harm S is if engaging with fictional-x_i *caused* S to harm another: for, by harming another, S creates the condition by which it is justifiable to harm S (Bloomfield, 2008). In short, if, on account of engaging with fictional-x_i, S harms another, then S harms himself, thereby producing (among other things, to be discussed below) a prudentially bad outcome. Thus:

- S is harmed *if*, on account of E (*qua* engaging with fictional-x_i), S harms another person, because harming another person makes things worse for S (thereby satisfying $HC_{suffering}$), owing to the fact that S is now in a position whereby some aspect of his liberty can justifiably be taken from him.[8]

On this account, the reason S harms himself is not because he engages with fictional-x_i per se but because he harms someone allegedly as a result of engaging with fictional-x_i. Fictional-x_i therefore causes him to harm someone else and, as a result, harm himself. More formally:

> HC_{causal} Engaging with (*qua* fictively imagining) fictional-x_i harms S *if* engaging with fictional-x_i causes P and, on account of P, S fares badly irrespective of temporal, counterfactual, or normative biological comparisons, or things are worse for S at t_2 compared to t_1 or than *if* P had not occurred.

Things are worse for S on account of P, and given our focus regarding what makes things worse for S, at least for the moment, is on the fact that S *harms another* and, in doing so, *harms S*, it is pertinent to ask: What are the ways in which engaging with fictional-x_i could cause S, on account of P, to harm another person and thereby harm himself? In other words, what constitutes tokens of P? In response to this question, consider the following supplementary harm condition (SHC):

> $SHC_{behavior}$ Engaging with (*qua* fictively imagining) fictional-x_i causes S to harm another person *if* engaging with fictional-x_i directly or indirectly causes S's behavior (P) to become more aggressive or otherwise antisocial, which includes a reduction in or an omission of behavior that would otherwise have helped prevent a separate harm from occurring and/or would otherwise have assisted someone suffering harm.

$SHC_{behavior}$ tells us something about how engaging with fictional-x_i could satisfy HC_{causal} and therefore what could be involved in harming another person and, on account of this (and in accordance with Mill's harm principle), what could be involved in harming oneself. The causally direct behavioral change mentioned in $SHC_{behavior}$ could arguably occur as a result of classical or operant condition (Skinner, 1938; Watson, 1913). *Pace* Watson and Skinner's respective brands of behaviorism, however, I am inclined to posit, as an intermediary step between engagement and behavior, some form of cognitive or conative influence (Pollock, 2006), thereby making engaging with fictional-x_i indirectly causally relevant to the harming of others and, therefore, oneself. To illuminate this intermediary step, consider two further SHCs:

> SHC_{belief} Engaging with (*qua* fictively imagining) fictional-x_i causes harm *if* doing so causes one to form a *belief* with content corresponding to the make-belief required to engage with fictional-x_i (e.g., the belief that sexually assaulting women is acceptable).

> $\mathrm{SHC}_{conation}$ Engaging with (*qua* fictively imagining) fictional-x_{i} causes harm *if* doing so causes one to develop the conation to engage in the type of behavior depicted by the fiction (e.g., engaging with depictions of fictional sexual assault on women causes one to cultivate the desire to assault women, sexually).

Neither SHC_{belief} nor $\mathrm{SHC}_{conation}$ has anything to say about *why* forming a belief or cultivating a desire of the kind described above is harmful to oneself or another. One speculative attempt at positioning one or more of the SHCs within an explanatory account of $\mathrm{SMC}_{behavior}$, and hence $\mathrm{HC}_{\mathrm{causal}}$, is outlined below. After discussing this, I will briefly touch on other ways in which SHC_{belief} and $\mathrm{SHC}_{conation}$ are harmful to oneself and others (although the reason for $\mathrm{SHC}_{\mathrm{conation}}$'s harmfulness will be taken up again in chapter 7).

A Speculative Account

When SHC_{belief} and $\mathrm{SHC}_{conation}$ are combined, it could be that the newly formed belief and conation cause congruent behavior (congruent, that is, with the aforementioned belief and conation), thereby satisfying $\mathrm{SHC}_{behavior}$ and, ultimately, $\mathrm{HC}_{\mathrm{causal}}$. One might surmise that, in tandem, they cause one to engage in sexually inappropriate behavior because one *believes* it is acceptable to behave this way and one *desires* to do so. Alternatively, in accordance with $\mathrm{SHC}_{conation}$, the newly acquired conation causes a change in behavior congruent with this conation, even in the absence of the belief described in SHC_{belief}. In other words, sexually inappropriate behavior is caused by one's desire to perform the behavior despite believing that it is unacceptable to do so or despite lacking the belief that it is acceptable to do so. Of course, whether this is the case, and therefore whether this speculative account gains purchase, is an empirical question. The extent to which there is support for the role of SHC_{belief} and $\mathrm{SHC}_{conation}$ will be taken up in the next chapter.

Harm through Omission, and Desensitization

At this point, it is worth noting that while $\mathrm{SHC}_{behavior}$ (and hence $\mathrm{HC}_{\mathrm{causal}}$) is satisfied *if* engaging with fictional-x_{i} causes conditions that themselves cause harm through *doing* (i.e., by producing conditions—whether conative or cognitive—that cause one to discriminate), $\mathrm{SHC}_{behavior}$ (and hence $\mathrm{HC}_{\mathrm{causal}}$) is also satisfied if fictional-x_{i} causes harm through omission (i.e., by producing conditions that cause one to *refrain* from certain behavior; see Foot, 1967; Feit, 2019, for a more recent discussion). Omission may occur on account of the acquisition of desires, or beliefs and desires, of the kind described by SHC_{belief} and $\mathrm{SHC}_{conation}$. I may, for example, fail to prevent A—a person who is nonwhite—from suffering harm if I hold racist beliefs and desire to see A harmed on account of these beliefs. But, equally, omission may occur on account of the fact that I do not perceive what is happening to A to be harmful. In the case of witnessing a man sexually harassing a women, I may not intervene because I believe that women "like it when men are forceful." In each of these examples, the constraint I show could be attributed to my racist or sexist views, but it need not be. Instead, after regular exposure to fictional-x_i, I may become *desensitized* to violence and/or other immorality, or become subject to what Kershnar (2005) calls nonautonomous changes, and so refrain from intervening for this reason. In recognition of this possibility, consider the following SHC:

> $\text{SHC}_{desensitized}$ Engaging with fictional-x_i causes harm *if* doing so causes one to become desensitized to the types of action and events that are depicted by the fictional immorality.

Continuing with the speculative account started earlier, in the case of $\text{SHC}_{desensitized}$, the fact that I am desensitized to depictions of fictional violence (for example) may cause me to behave in a manner that others perceive as aggressive because I am desensitized to the inappropriateness of my aggressive way of resolving a dispute. Or, in the case of refraining from intervening to prevent harm to another, because I do not perceive the situation to be one in which the person needs assistance (recall the sexual harassment example).[9] Owing to my engagement with depictions of fictional violence (e.g., torture and murder), I may behave aggressively, or in an otherwise antisocial manner, including refraining from assisting someone suffering harm (all in accordance with $\text{SMC}_{behavior}$), because I have become desensitized to violence and its harmful consequences. Moreover, desensitization may affect my beliefs (e.g., I may come to believe that "*that*'s not violent") and my desires (I may grow to "like doing this"—not least because I no longer respond to the depiction in a viscerally negative way (e.g., disgust), and so it is not incompatible with SHC_{belief} and $\text{SHC}_{\text{conation}}$.

One must be cautious in the case of $\text{SHC}_{\text{desensitized}}$, however, because, even if desensitization occurs, it is not inevitable that this will lead to a change of behavior, directly, or even indirectly, in virtue of a change in belief and connation. After all, one could speculate that police officers (and such like) who are regularly exposed to violence are likely to exhibit signs of desensitization without further suggesting that they are more likely to be aggressive or less likely to help those in need. Indeed, Hinte (1971) asks us to consider whether techniques for desensitization ought to be part of police recruitment training. Given this, it is perhaps appropriate to postpone further talk about desensitization until the section "The Risk of Increased Harm," when discussing the risk of increased harm.

It may also be the case that engaging with fictional-x_i does not necessitate that one must acquire the *corresponding* belief or desire, as expressed by SHC_{belief} and $SHC_{conation}$; it may be, instead, that one acquires more generic violent beliefs and desires that, in turn, motivate more general acts of aggression/antisocial behavior. So, after engaging with depictions of fictional torture and murder, I may behave aggressively, or in an otherwise antisocial manner, without attempting to murder or torture anyone or without holding the belief that this is acceptable or possessing the desire to do it. Again, the extent to which my conjecture finds empirical support will be discussed in the next chapter.

Endorsement Revisited

In chapter 4, I argued for the immorality of a fiction that *endorses* an immoral worldview, and therefore for the immorality of any individual depiction of fictional-x_i contained therein that is congruent with the immoral worldview endorsed. If, through engagement with fictional-x_i, irrespective of whether the depiction is congruent with a fiction that is endorsing an immoral worldview, and therefore irrespective of whether the meaning conveyed by the depiction is immoral, S starts to endorse that which fictional-x_i depicts (i.e., S comes to believe and support—tacitly or otherwise—the immoral worldview that sexual violence toward women is acceptable), then, in accordance with SHC_{belief}, and on account of forming the belief, S is worse off. This is because S is now of the view that that which is immoral is acceptable to perform, and such private endorsement is immoral. Even if S never expressed this belief, or acted in a manner congruent with it, what they now believe is immoral and, as a result, satisfies HC_{causal}. This is because things are *now* worse for S than they were before (temporal comparison), on account of P (namely, the formation of the immoral belief), or than they would have been if P had not occurred (counterfactual comparison). It is also important to note that holding an immoral worldview about the acceptability of sexual violence toward women makes things worse for them, and therefore harms others (*qua* women) as well as oneself, because, at the very least, women are *regarded* as persons toward whom violence is acceptable, even if one never inflicts such violence on them, or attempts to do so (Bartel & Cremaldi, 2018). Moreover, irrespective of belief, if, as a result of engaging with fictional-x_i, S acquires or develops an existing liking for the sorts of immoral acts depicted by the fiction, then this acquired or cultivated predilection for certain sorts of immorality is potentially problematic for reasons

I will discuss below and return to in chapter 7 when examining the morality of enjoying fictional immorality.

HARMING OTHERS

Discussion on the (potential) harmful effects of engaging with fictional-x_i has thus far focused on the harm caused to oneself. Where others are harmed, on account of one's engagement with depictions of fictional immorality, this harm has been referenced only insofar as it acts as a means by which one harms oneself. But the potential harm done to others, on account of S's engagement with fictional-x_i, may strike many readers as of equal if not greater concern. Engaging with fictional-x_i cannot, of course, directly harm those who do not engage with it.[10] If it does harm others, then, as discussed, this can only be achieved indirectly on account of some change to those who do engage with fictional-x_i. For one's engagement with fictional-x_i to harm another, it would have to cause this harm through its effect on the one who is engaging with the fiction. This means it would have to *cause* them to do something that, in turn, causes the other person to be in some sense worse off. Should this occur, then, as discussed, where S makes another worse off, on account of engaging with fictional-x_i (say, in accordance with $\text{SHC}_{\text{behavior}}$ which in turn may be further supplemented within the explanatory account by $\text{SHC}_{\textit{belief}}$ and $\text{SHC}_{\textit{conation}}$), then S has harmed another and, in doing so, harmed himself. Whether any of this transpires, of course, rests on a series of causal relationships: namely, S's engagement with fictional-x_i causing a change in S, which then causes a change to the other person (making things worse for them, thereby making things worse for S). But what if instead of a causal relationship between S's engagement with fictional-x_i and a change in S, engaging with fictional-x_i simply increased the *likelihood* of a change in S, thereby only increasing the likelihood that S would harm another on account of engaging with fictional-x_i?

Consider what this HC might look like:

> HC_{risk} Engaging with fictional-x_i harms S *if* engaging with fictional-x_i *increases the likelihood* of P and, on account of P, S fares badly irrespective of temporal, counterfactual, or normative biological comparisons, or things are worse for S at t_2 compared to t_1 or than *if* P had not occurred.

The SHCs already presented can be easily adapted to fit HC_{risk}. All that needs to change is that instead of *causing* P (where P is a change in behavior, as stipulated by $\text{SHC}_{\textit{behavior}}$, or a change in belief and/or conation and/or desensitization as stipulated by the other SHCs, respectively), engaging with fictional-x_i simply makes P more likely. Thus making it more likely *but not inevitable* that engaging with fictional-x_i will harm others and therefore oneself.

Even if we allow that a change in S (caused by engaging with fictional-x_i) will *cause* a change in another person, thereby making things worse for them (and subsequently worse for S), what should our moral position be, regarding engaging with fictional-x_i, if it transpires that such engagement, while increasing the likelihood of causing a change in S that would result in S harming another (and subsequently himself), is not inevitable. In other words, if it turns out that S's engagement with fictional-x_i is, at the very least, not sufficient for S to cause harm to another, but is sufficient to increase the likelihood of harm occurring, then how should we classify what S is doing and what should our moral position be with regard to our engagement with fictional-x_i *qua* a risk (of harm) enhancing activity? Let us consider the first part of this question first.

THE RISK OF INCREASED HARM

It is my contention that S *is* worse off on account of the fact that, through his action, he is now in a position where the likelihood of harm occurring has increased. Consider S at t_1 (low likelihood of harm, let us allow) compared to t_2 (increased likelihood of harm). S is worse off at t_2 (after the risk-increasing event) compared to t_1 (before the event) or compared to the counterfactual condition in which the risk-increasing event does not occur. Consequently, his action fits one of the conditions for harm we have been discussing: namely, $HC_{suffering}$ (recall: S is harmed *if*, on account of E, S fares badly irrespective of temporal, counterfactual, or normative biological comparisons, or things are worse for S at t_2 compared to t_1 or than *if* E had not occurred). So, even though it is not inevitable that S harms another and, in doing so, harms himself, it is nevertheless the case that engaging with fictional-x_i, where so doing increases the likelihood of harm occurring, is harmful to S, *even if S does not harm anyone else* on account of engaging with fictional-x_i. Put differently, engaging with fictional-x_i is harmful even if it does not *cause* or otherwise contribute to S harming another and, as a consequence, himself. To be clear, such harm does not have to be salient. I am not suggesting that S must suffer in any experiential sense, only that S is made worse off on account of the increased risk of harm.

In support of the last point, consider the distinction McCormick (2001) makes between:

(i) A *dangerous act*, which is an act that directly increases the risk of harm to self or others (e.g., engaging in a knife-throwing act).

(ii) A *harmful act*, which is an act that results in direct injury or damage to self or others (e.g., hitting one's assistant with a knife or stabbing oneself in the foot with it).

(iii) A *risk-increasing act*, which is an act that increases the person's chances of committing a dangerous or harmful act (e.g., drinking alcohol while deliberating over whether to perform a knife-throwing act).

McCormick introduces this distinction when discussing the potential harm caused by playing "violent" single-player video games. I shall retain McCormick's focus on video games, temporarily, and for convenience only, but with the understanding that the distinction he makes can be applied to depictions of fictional immorality across media. For McCormick, engaging in video game violence is not a dangerous or a harmful act as defined by (i) and (ii) because one does not physically harm, nor can one run the risk of physically harming, directly, either oneself or another as a result of what is done *within the game*. Engaging in video game violence could, however, be construed as a risk-increasing act: for although no harm is *directly* incurred through engaging in virtual violence (at no point am I ever actually assaulted or killed within the video game, nor do I actually assault or kill anyone), such activity does (potentially, and allegedly) increase the risk of engaging in the sorts of dangerous activities that themselves run the risk of directly causing harm to oneself or others.

To illustrate: If I participated in an unauthorized car race through the populated streets of my hometown and did so without striking a pedestrian or otherwise causing harm and injury to anyone, then while I may not have performed a harmful act, I nevertheless acted in a manner that could have caused injury and perhaps even permanent damage to others, including myself (insofar as I acted in a way that increased the likelihood of this happening). What I did, then, was to engage in a dangerous act. By contrast, playing a video game, in which I engage in a virtual car race through populated streets via my avatar (against a computer-generated opponent: e.g., *Carmageddon*), or watching a film featuring a similar race (e.g., *Death Race 2000*), is not, in and of itself, dangerous in the manner described by McCormick. This is because no one risks being directly harmed on account of my engagement with the depiction. Engaging with fictional immorality in this way is, however, an activity that (allegedly) increases one's risk of doing something that is dangerous (i.e., critics fear that playing such a game increases the likelihood that one will drive at faster speeds for real, and therefore more recklessly). Given this possibility, it is worth considering whether the following IC should apply:

> IC_2 Engaging with fictional-x_i is immoral *if* it *risks* (*qua* increases the likelihood of) causing harm.

IC_2 is grounded on the precautionary principle that states: "Where an activity raises threats of harm to the environment or human health, precautionary measures should be taken even if some cause and effect relationships are not fully established scientifically" (Bradley, 2012, p. 390; see also Powell, 2010, for a critical examination of the principle). A similar precautionary approach is advocated by Singer (2007) when discussing video game "violence":

> Manufacturers fall back on the simplistic assertion that there is no scientific proof that violent video games lead to violent acts. But sometimes we cannot wait for proof. This seems to be one of those cases: The risks are great and outweigh whatever benefits violent video games may have. The evidence may not be conclusive, but it is too strong to be ignored any longer. (p. 1)

Likewise, for Waddington (2007), potential risk should not be ignored: for, as he acknowledges, potential risk—the risk of a risk—is still a risk. Thus, if we accept Waddington and Singer's advice, we should include even the *potential* for increased risk within our assessment of the morality of engaging with depictions of fictional immorality, regardless of whether empirical proof is presently forthcoming to indicate that this potential is being realized on a consistent basis.

In short, then, following McCormick's (2001) distinction, event E is not dangerous because it does not increase the likelihood of harm occurring directly (as would be the case if one were to drive at high speed in a populated area). Instead, it is a risk-increasing event. It increases the likelihood that one will engage in a dangerous act (which increases the likelihood of harm occurring directly) or engage in a harmful act itself. Again, using the car example, it increases the likelihood that one will engage directly in dangerous behavior, say, by changing one's attitude to safe driving and one's desire to drive quickly. What is potentially ambiguous about what is being discussed, however, is the claim that participating in a risk-increasing event makes one worse off and, because of this, is harmful. It is ambiguous because McCormick distinguishes between a risk-increasing event and a harmful act; yet, it would appear that engaging in something that is risk increasing is a harmful act.

A way to resolve this potential ambiguity is to make clear that, for McCormick, a harmful act is one that results in *direct injury or damage to self or other*, whereas the harmfulness of a risk-increasing act/event is that it makes one worse off insofar as it places one at a disadvantage. The disadvantage is that one is now, on account of one's action, in a position whereby one is at increased risk of doing harm *qua* causing injury or damage to oneself or another. This increased risk, while harmful insofar as one is now worse off, does not itself cause damage or injury; rather, it may cause a change in oneself such that it makes injury or damage to self or another more likely.

Should IC_2 be upheld in cases where one risks harm to oneself only? In response to this question, I favor the interpretation of Mill's harm principle provided by Saunders (2016), which restricts one's liberty to harm to cases where there is consent and therefore voluntariness. As Saunders illustrates:

> Suppose that I consent to a boxing match and, as a result, suffer a broken nose. It may seem odd to call this a consensual harm, since I never consented to having my nose broken. However, perhaps we may say that I consented to the risk of a broken nose, *knowing that this may occur* when boxing. (2016, p. 1011; emphasis added)

Likewise, where S voluntarily and competently consents to insert variously sized metal pins into his limbs and torso as part of a body art exhibit and, by doing so, inadvertently inflicts irreparable damage on himself, he did not consent to *this* harm happening (only a *pro tanto* harm),[11] although he did consent to the *risk* of harming himself in this way. *Mutandis mutatis*, the same can be said of Harry Houdini when he entertained the crowds with his "death-defying" escapes, or of the lone sailor who seeks to break the record for the fastness or the youngest transatlantic crossing, or achieving this in the smallest craft, and so on. In each case, as with the boxer, what they consent to, perhaps tacitly, is to risk having irreparable damage inflicted on them; and, in each case, they are (and, in the case of Houdini, were) at liberty to do this.

For Mill, of course, such liberty is grounded on utilitarian principles. In each of these examples, one could argue that the risk-increasing action each agent has consented to undertake is considered, by the agent, to be outweighed by the anticipated benefit. Or, in accordance with Kant's deontological position, their consent-based liberty accords with their status as autonomous rational agents. By contrast, and again in accordance with Saunder's interpretation, one should be prevented from engaging in activities that risk harming oneself if one is under mental duress: for, under mental duress, one is not able to consent appropriately (in an informed way) and is therefore not acting in a suitably voluntary manner.[12]

Returning to $SHC_{\text{desensitized}}$, while there is a case to be made in favor of the harm of acquiring certain beliefs and connation as a result of engaging with fictional-x_i (as discussed)—that these can be, in a qualified sense, harmful and hence immoral irrespective of causing a change in behavior—the harm brought on by desensitization is less clear-cut. In other words, it is not apparent how becoming desensitized to depictions of fictional immorality and what these are depictions of is harmful in any way other than because desensitization *increases the risk* of altering one's beliefs and/or connation and/or behavior. It may be, instead, that desensitization, far from being a problem, helps the audience overcome their visceral responses to violence:

responses that would otherwise detract from their ability to appreciate the aesthetic quality of the depiction, for example, by forcing them to look away or otherwise cease to engage (Symonds, 2008). Whether desensitization does lead to any of these changes is, of course, an empirical question, which will be taken up in the next chapter.

What about the increased risk of harming another through one's actions? It is one thing to have to consent to the risk of harm regarding oneself, but how is it possible for those members of society not directly involved with the risk-increasing fiction to give their consent? For IC_2 to be upheld in the case of the risk of harm to others, much rests on the nature and amount of evidence supporting this increased risk: How much risk? From whom to whom? And how much support is there for this relationship? Again, this will be taken up in the next chapter.

CONCLUSION

In this chapter, I presented a case for the immorality of harm. I also set out what I take to be a number of sufficient conditions for the occurrence of harm, as well as the means by which these conditions might be satisfied. Where there is a lack of a direct causal relationship between engaging with fictional-x_i and harm, I discussed the possibility that engaging with fictional immorality might nevertheless *increase the risk* of harm occurring. In the case of increased risk solely to oneself, I argued for consent-based liberty to perform the risk-increasing activity. As for an increase in the risk of harm to others (who are not in a position to give their consent), permitting the agent to act depends on other factors, not least of which is evidence supporting the claim that there is an increased risk of harm to others (and likely oneself, for reasons discussed). What I have intentionally omitted from the discussion thus far, however, are the findings of research looking at the effects of engaging with depictions of fictional violence or other immoral acts across different media. What will be considered in the next chapter, then, is the evidence for and against the claim that engaging with depictions of fictional immorality either causes harm or increases the likelihood of harm occurring to oneself and others.

NOTES

1. For a small sample of media articles on the topic, particularly video game "violence," see https://edition.cnn.com/2016/07/25/health/video-games-and-violence/index.html; https://www.cbsnews.com/video/video-game-violence-linked-to-bad

-behavior-study-says/; https://www.scientificamerican.com/article/yes-violent-video-games-trigger-aggression-but-debate-lingers/ (all accessed June 21, 2019).

2. Thomson (2011) makes a similar point using an example of a firefighter who, in order to rescue someone from a burning building, has to break their arm.

3. To be worse off is not, of course, simply to be in deficit. If I *willingly* donate time, money, other resources to charity, for example, I am not worse off for doing so.

4. According to Harris (1998), to say that "things are worse for S" is effectively to say that S is in a state that a rational person would not want to be in. A problem with the claim to rational preference, however, is that it would seem to exclude non-human animals and any individual incapable of being rational (including babies/small children) from suffering harm. For Kahane and Savulescu (2012), in contrast, one is worse off if one is placed in a state or condition that falls below what is statistically normal. This approach has its own problems. Plague victims would not be harmed by contracting the plague if contracting the plague was not below what was statistically normal during the Black Death (see Purshouse, 2016).

5. By using this example, I am claiming only that, in having one sense fewer (e.g., sight) than is standard, S is worse off than if all of his senses were intact. This being the case, however, does not necessitate or even suggest that S's overall quality of life is inferior.

6. The use of the term "biological" is for nomological convenience; it is not a metaphysical necessity.

7. Depriving someone of their liberty is one of Gert's five basic harms (Gert, 2004).

8. Point of clarification: Restricting someone's liberty can take many forms and occur to varying degrees. It does not entail imprisonment. One way, in the context of fictional immorality, may be to deprive S of his liberty to engage with fictional immorality.

9. Stoner (2020) defends "gorefest" horror fiction against the argument from *reactive attitudes*, which holds that it is wrong to watch such graphically violent films because the extreme depictions of blood and gore they contain threaten to damage our ability to react to real-life examples of suffering with appropriate levels of compassion. Stoner argues that there is no compelling evidence that depictions of graphic violence presented in other media have this effect and neither is there evidence or argument to support the view that gorefests are somehow a unique case.

10. It could be argued that *allowing* anyone to engage with fictional-x_i harms society as a whole because of what it says about us—that we are the sort of society that allows people to engage with depictions of fictional immorality knowing that they cause harm. Such an objection only works, of course, if engaging with fictional-x_i is harmful to begin with or is immoral based on some other criterion. If the latter, this takes us away from the issue of harm we are discussing here. I will, however, return to this point in the next chapter.

11. The same might be said of the boxer: that they consider the harm that will no doubt be inflicted on them as part of the boxing match to be a *pro tanto* harm, given that they intend for the overall outcome to be beneficial, say, by winning the bout and perhaps even becoming world champion.

12. To say that causing harm, even to oneself, in the absence of competent consent is morally wrong is not to say (insofar as it does not necessitate) that the person responsible for causing the harm is immoral or morally responsible for their action. A distinction can and should be drawn between the immorality of the act itself (say, inflicting damage on oneself) and the moral culpability of the one inflicting the damage.

Chapter 6

Is There Evidence of Harm?

INTRODUCTION

In the previous chapter, I presented ways in which engaging with fictional-x_i might harm another person and subsequently oneself. Suggestions were based both on a causal connection between the engagement and behavior—whether direct or indirect—and the increased likelihood of harm, rather than its inevitability. Importantly, I deliberately omitted discussion on the findings of empirical research that have looked at the effects of "media violence" on behavior, cognition, and affect. These findings are the focus of this chapter. Before considering them, however, I present historical examples of concern over fiction, particularly fictional depictions of immorality, and also discuss fictional immorality's relationship to obscenity and harm. After that, I present what amount to rather mixed findings on the impact of media violence on aggressive behavior, affect, and cognition across a number of different media: with some studies supporting a negative impact and others refuting it. The contradictory nature of these research findings results in a lack of consensus on the impact of violent media, which, in turn, reveals a lack of universal support, or anything close to this, for the claim that engaging with depictions of violent media causes harm or even increases the risk of harm occurring.

In the absence of consistent evidence, it is difficult to *justify* adopting a nomothetic approach to the matter of harm. Therefore, even if one were to support either of the immorality conditions presented in chapter 5, *in principle*—although "in principle" support is perhaps more straightforward in the case of IC_1 than IC_2—there is currently no consistency in research evidence and, subsequently, no consensus on which immorality condition, if any, finds empirical support in the context of fictional immorality.

HISTORICAL CONCERNS

Concern over the alleged increased risk of harm brought about by depictions of fictional immorality is nothing new. In *The Republic*, not only is Plato critical of art because, as he sees it, it is incapable of guiding us to the truth of the forms or ideas (and, instead, moves us away from these toward illusion[1]); importantly, owing to art's capacity to elicit strong emotional responses, it should be treated with caution, even when relegated to mere entertainment. The poet who elicits laughter through mockery or rejects the triumph of nobility over adversity is, according to Plato, worthy of censorship, at least as far as a society's youth is concerned: for exposing young minds to such fiction is harmful.

> And shall we just carelessly allow children to hear any casual tales which may be devised by casual persons, and to receive into their minds ideas for the most part the very opposite of those which we should wish them to have when they are grown up?
>
> We cannot.
>
> Then the first thing will be to establish a censorship of the writers of fiction, and let the censors receive any tale of fiction which is good, and reject the bad; and we will desire mothers and nurses to tell their children the authorised ones only. Let them fashion the mind with such tales, even more fondly than they mould the body with their hands; but most of those which are now in use must be discarded.
>
> (Plato, *The Republic*, Book II, 377b–377c)

For Plato, bad fiction—in the form of poems or plays—depicts those of low moral character making bad choices, seemingly for our entertainment: eliciting from us a kind of Schadenfreude. Indulging such emotion cultivates poor values and subsequently citizens of low character, which should not be encouraged. Hence, such emotion-eliciting fiction should be purged through censorship. Importantly, though, while Aristotle shares Plato's views on the importance of good character, in contrast, he does not consider art and the powerful emotions it can elicit to be a reason to condemn it (Eden, 1982); rather, for Aristotle, art, and hence fiction, provides the means for us to *purify* our emotional repertoire through *catharsis*, and so refine our emotions to the point where our habituated responses to emotionally arousing events become measured: situated in the middle (or as the mean) of two extremes or vices (e.g., too much fear results in cowardice, whereas not enough can lead to recklessness). As Gentile (2013) notes: “[The] ability to feel the extremes and view the consequences acted out by an actor or poet allows the viewer to begin to find the mean between the extremes” (p. 497).

The kinds of emotion Aristotle has in mind are predominantly fear and pity. In order for these to be tempered, for the sake of building good character, they need to be aroused by artistic (fictional) depictions of tragedy, designed with this cathartic aim in mind (Daniels & Scully, 1992). Yet, for Gentile (2013), much of today’s violent fiction or even comedy is not formatted to arouse the kinds of emotion Aristotle had in mind.

For illustrative purposes only, consider Di Muzio’s (2006) objection to certain sorts of horror film—namely slasher films or what, today, we might call torture porn (Edelstein, 2006). According to Di Muzio, these types of film focus on terror, torture, and mutilation (Johnston, 1995) and “are *devoted primarily or solely* to representing violence and death” (Di Muzio, 2006, p. 281; italics in original), to the extent that they “[make] the point of having no moral point” (*ibid.* p. 290). In order to apply Aristotle’s approach to fiction to a modern-day example, let us accept Di Muzio’s interpretation, although I must stress that it is not without its dissenters (see below). Under such an interpretation, slasher films/torture porn have diminished plot, character development, thought, and diction (each of which Aristotle values, in that order, for good drama to occur) and, instead, overemphasize spectacle (Aristotle’s least valued component of good drama)—but not just spectacle, spectacle of the wrong sort. Within such films, Di Muzio continues, visually graphic gore and violence are presented as a feature of the entertainment and in a way that does not appear to provide an opportunity for us to reflect on the violence graphically depicted. For the sake of balance, however, it is important to note that Kreider (2008), in contrast to Di Muzio, argues that depictions of violence in slasher films are not typically an end in themselves and that often reasons for the violence are contained within the narrative, usually waiting to be uncovered as the story unfolds.

Leaving this argument aside, if films of the kind Di Muzio describes were actually to exist—and there is no reason for them not to, in principle—then Aristotle would have this to say of them:

> Tragic fear and pity may be aroused by the spectacle; but they may be aroused by the very structure and incidents of the play—which is the better way and shows the better poet. The plot in fact should be so framed that, even without seeing the things take place, he who simply hears the account of them should be filled with horror and pity at the incidents; which is just the effect the mere recital of the story in Oedipus would have on one. To produce this same effect by means of Spectacle is less artistic, and requires extraneous aid. Those, however, who make use of the Spectacle to put before us that which is merely

> monstrous and not productive of fear, are wholly out of touch with Tragedy. (Aristotle, Poetics, 14, 1453b1–6)

For Aristotle, the effect of catharsis—to purify one's emotional response—is arrived at preferably through well-crafted plot and diction. It can be achieved through spectacle, but, according to Aristotle, this is of less artistic merit. Aristotle would therefore be critical of any drama (not just one from within the slasher-horror genre, if Di Muzio is to be believed) that emphasized spectacle over plotline and characterization (see Haynes, 2016), particularly where this promotes monstrous spectacle. That is, spectacle at the expense of promoting catharsis *qua* the purification of pity and fear.

Concerns over the effects of engaging with fiction (in general) not just spectacle, and not just depictions of immorality, have been raised over the centuries. In 1750, for example, Samuel Johnson wrote, in *The Rambler*, about the potential dangers of realism in fiction owing to its power, as he saw it, to influence the reader, particularly the impressionable youth. At around the same time, novels such as *Thérèse Philosophe* (Therese the Philosopher) and *Fanny Hill: Memoirs of a Woman of Pleasure* were courting controversy, owing to their depictions of (alleged) immorality, as was *L'École des Filles* (The School for Girls) about a century earlier (1665); and let us not forget the once-reviled author Donatien Alphonse François, more commonly known as the *Marquis de Sade*, whose works have only recently begun to receive literary accolades.[2] (For an interesting discussion on reactions to immorality in Victorian fiction, see also Garrison, 1976.)

Literature of the kind just mentioned was often censored or banned at the time of publication and, in some cases, for years afterwards (e.g., in the UK, the novel *Fanny Hill* was banned until the 1960s). A common charge against works such as these was that the content (at least, in part) was indecent and/or obscene. Obscenity is a legal term, not a moral one, although, in everyday parlance, it is often used as an expression of moral disapproval. Its relevance to this chapter is therefore that it is often associated with, even if not intended as a direct measure of, harm (e.g., the United States), although, in the UK, the legal definition of obscenity makes mention of its required moral corruption, and hence harm.

OBSCENITY AND ITS ASSOCIATION WITH HARM

In 2001, when discussing whether the U.S. obscenity test[3] should be applied to the content of "violent" video games (case: *American Amusement Machine Association v Kendrick*[4]), the presiding judge, Posner, reasoned thus:

> The main reason for the proscription of obscenity "is not that it is harmful, but that it is offensive." Obscenity is regulated because people find it "disgusting, embarrassing, degrading, disturbing, outrageous, and insulting," not because it is "believed to inflict . . . harm." (Taken from Stone, 2007, p. 1858)

In contrast to the U.S. measure, within the U.K. definition (stemming from the 1959 *Obscene Publications Act*), rather than the measure of obscenity being rooted in some form of offense principle, classification is based on whether the material is likely to deprave or corrupt those who have access to it. Something is obscene:

> If its effect or . . . the effect of any one of its items is, if taken as a whole, such as to tend to deprave and corrupt persons who are likely, having regard to all the relevant circumstances, to read, see or hear the matter contained or embodied in it. (Section 1:1)

With the U.K. definition, what is considered obscene is couched in social pathology such that there would be a tendency toward "moral and physical harm caused to vulnerable persons by exposure to obscene writings and images" (Hunter et al., 1993, p. 138). In addition, and again showing a departure from mere offense toward social pathology, one's reaction to obscenity must be more than disgust, even when this involves physical revulsion and/or behavioral aversion: one must become (or be in danger of becoming) morally corrupted. Section 1 of the Obscene Publications Act thus tries to distinguish between what merely offends communal standards of acceptability and what is socially harmful.[5] Perhaps with this in mind, Stone (2007), in relation to the U.S. obscenity test, argues that although offensiveness may well be part of the *definition* of what it means to be legally obscene, this should not rule out the possibility, even the likelihood, that regulation is enforced because obscenity causes or is believed to cause harm.[6]

Obscenity and Moral Harm

Kieran (2002) defines obscene representations as those that solicit from us cognitive-affective responses toward objects/events that are morally prohibited, to the extent that we are commended "to delight in them" (p. 41). What we are prescribed to delight in, Kieran informs us, are morally prohibited sexual acts, or the infliction of pain or suffering or even death of another, either by one's own hand or vicariously. To commend us to delight in that which is already demarcated as morally prohibited, even taboo, is taken to be not only a measure of the obscene, *and* the implied intention of the creator of the material, but also a consequence of it for us, at least if the exposure to the obscene material is prolonged. Thus, it may be argued that obscenity is morally wrong because it morally corrupts; and by “morally corrupts,” we mean that it not only *commends* us to delight in the morally prohibited but also, and importantly, *causes* us to do so. The association of obscenity with moral corruption is further implied by Koppelman (2005), when commenting on U.S. legislation:

> Material can be obscene even if it has no likelihood of inciting anyone to unlawful conduct, and even if no unwilling viewer is ever likely to see and thereby be offended by it. Obscenity law aims at preventing the formation of certain thoughts—typically, erotic ones—in the minds of willing viewers. (p. 1637)

Even in the absence of illegal activity, and even in the absence of offense, Koppelman claims that something can still be deemed obscene if it leads to the formation of certain thoughts—most likely about morally prohibited acts or the desire to engage in or witness such acts—even in the minds of the willing. Obscenity is therefore judged to be detrimental to our psychological well-being, even if “psychological well-being” is restricted to cognitive-affective states, thereby excluding behavior. In part at least, the aim of obscenity legislation is to prevent the formation of such cognitive-affective states, perhaps because there is an implicit assumption that their formation will lead to, or greatly increase the risk of, violating morally proscribed, most likely illegal, behaviors (recall discussion on the supplementary harm conditions in the previous chapter: $SHC_{behavior}$, SHC_{belief}, $SHC_{conation}$).

Leaving aside the question of obscenity (see Young & Whitty, 2012, for a more detailed discussion), our concern for the remainder of this chapter is not the offense caused by fictional-x_i but the likelihood of *ensuing harm*, and not just to the one who engages with the fiction but to those whom they might encounter. Unlike offense, then, which has a subjective quality, harm seems more objective, based on “some set of observables” (White, 2006, p. 119).

MEDIA VIOLENCE AND EVIDENCE OF HARM

In this section, I discuss general research findings, particularly from meta-analyses, on the effects of media violence on behavior. It is not my intention to assess the merits or failings of individual studies; rather, I will discuss general conclusions, arrived at through the culmination of numerous studies carried out over many years and using various methodologies. Initially, discussion will include only the most popular and subsequently most researched media: films, television, and video games (although I will touch on comic books and cartoons, here, also). After that, I will selectively examine other media, namely song lyrics and novels (section "Other Fictional Spaces"). I will also adopt the definition of media violence used in the 1998 National Television Violence Study (see below) as it contains key features that are typically examined in research on media violence.

According to the 1998 study, media violence amounts to:

> Any overt depiction of a credible threat of physical force or the actual use of such force *intended* to physically harm an animate being or group of beings. Violence also includes certain depictions of physically harmful consequences against an animate being or group that occurs as a result of unseen violent means. Thus there are three primary types of violent depictions: credible threats, behavioral acts, and harmful consequences. (p. 41; cited in Lindsay et al., 2014, p. 7; emphasis added)

The definition does not restrict depictions of violence to acts between humans. As Busching et al. (2016) point out, this allows for violent depictions to include the aggressive antics of nonhuman animated characters such as Wile E Coyote and the Road Runner, Mutant Teenage Ninja Turtles, or Transformers. Nor is blood and gore part of the definition. Violence can therefore be depicted in the absence of this and still be classified as violence. Another important point to note is that research on media violence often does not specify whether the depictions examined constitute what I have been calling pure fictional immorality. While reference to films, television, and video games within research literature will often imply fictional content of the kind that satisfies my definition, this cannot be guaranteed. With this caveat in place, it is still worth considering what the evidence on media violence has to say about potential harm.

General Media Violence

For decades, media depictions of violence have increased in frequency and graphic realism (Barranco et al., 2017; Bushman et al., 2013; Lyndsay et al., 2014; Thompson, 2004).[7] While this comment may strike the reader as rather unremarkable, and may even chime with their own understanding and experience of watching films and television dramas, or when playing "violent" video games, what is contested is the role media violence plays in generating violent or otherwise antisocial behavior. A common retort directed at those who criticize the (ever-increasing) amount of media violence available is that media violence simply *reflects* the violence within society, rather than causes it. As such, it is the violence that already exists within society that requires these media to portray increasing amounts of violence as a way of maintaining authenticity. This claim is not true, however. Media violence, in fact, often *exaggerates* societal violence (Lyndsay et al., 2014). To illustrate, Mount (2008) reports that in the television drama *Lewis*, the eponymous Oxford-based detective solved eight murders in a month. As the population of Oxford (in the UK) is approximately 150,000, and if media violence is simply a reflection of societal violence, then Oxford would have a murder rate eleven times higher than New York. The number of actual murders in Oxford that same year was one. This example of exaggeration is not unique. Lamont (2018) likewise reports that the murder rate on the remote Scottish island of Shetland, at least based on the murders that have occurred within the last four series of the TV drama *Sheltand*, is 68.2 per 100,000. This would make Shetland the eleventh deadliest place to live in the world. In actual fact, there have been two reported murders on Shetland in the past fifty years.

If media violence exaggerates rather than accurately reflects societal violence, and the amount of media violence is increasing, then what impact is this having on societal violence and antisocial behavior more generally? Numerous studies purport to show a connection between media violence and increases in aggressive behavior (Anderson et al., 2003, 2017; Atkin, 1983; Bushman et al., 2015; Goranson, 1970; Huesmann & Taylor, 2006; Strasburger & Wilson, 2014). Busching et al.'s (2016) summary (below) is typical of those supporting the view that consuming media violence leads to increased aggression.

> Media violence exposure leads to higher levels of aggression. This finding is supported by a large number of studies utilizing a wide variety of methodological approaches. The connection between media violence and aggressive behavior is explained by many different mediational processes. Violent media exposure can change what is considered socially acceptable, how the environment is perceived, and how we feel about violence. All these processes combined lead a [*sic*] higher levels of aggressive behavior. (pp. 13–14)

Busching et al. also add that there is virtually no evidence for the opposite argument: that violent media help people "let off steam," thereby reducing aggressive urges through cathartic purging[8] (see also Gentile, 2013; Huesmann et al., 2013; Strasburger & Wilson, 2014). Moreover, in addition to increased aggression, they also note that "the effects of violent media are not limited to aggressive behavior. There is also substantial evidence linking violent media exposure to problems with attention, impulsivity, and executive functioning as well as reductions in empathy and prosocial behavior" (Busching et al., 2016, p. 14), and even, more recently, ethical decision making in business (Gubler et al., 2018).

A subset of the "violent media" research has focused on cartoons. After all, in many cartoons, including those written for children, violence is an integral part of the content. One has only to watch classic cartoons like *Tom and Jerry* or the *Roadrunner*, which is endlessly pitted against *Wile E Coyote*, to find support for this claim. Indeed, *Itchy and Scratchy*—a cartoon within the cartoon *The Simpsons*—with its extreme violence, can be said to parody *Tom and Jerry*.

Kirsh (2006) found that comedic violence within cartoons, of the kind found in the likes of *Tom and Jerry*, was associated with lower levels of aggression, compared to noncomedic cartoon violence. The comedic element, Kirsh suggests, acts to trivialize and/or camouflage the violence. This would appear to negate aggressive imitation compared to noncomedic forms of violent cartoons. In earlier work, albeit in relation to illustrated rather than animated comic violence, Kirsh and Olczak (2002b) found that reading extremely violent comic books had a negative effect on the processing of ambiguous situations within a fictional narrative compared to reading nonviolent comics (see also Kirsh & Olczak, 2000, 2002a). Similarly, Coyne et al. (2015) found increased aggression in adolescents who read manga graphic novels (which often contain violent fictional imagery) compared to nonmanga readers.

With such evidential support available across a range of media, why is the issue of media violence and societal violence contentious? Christopher Ferguson—who has consistently cautioned us against drawing too strong a conclusion about the causal relationship between media violence and societal violence, often arguing for no causal connection, at least based on current evidence—cites publication biases (the tendances for research showing a connection between media and societal violence to be published) and methodological problems that inflate effect sizes (Ferguson & Kilburn, 2009). Ferguson is not alone in his criticism of research allegedly demonstrating the negative consequences of media violence. Freedman (1984) claims no causal connection and only weak correlations (see also Freedman, 2002). Likewise, Savage (2004) reports no consistent evidence to support a relationship between violent media and criminal behavior (see also Ferguson, 2015, who, after examining rates of media violence and homicide in the United States, claims no consistent relationship). Following an extensive review of published research, Gunter (2008) concludes:

> Evidence for the effects of media violence on real-world violence has provided indications of effects but cannot be accepted as having presented a conclusive case. Each of the sources of this evidence is sufficiently questionable to limit its application. . . . Moreover, in relation to media events, convincing alternative explanations for increases in levels of social violence have been put forward. . . . Finally, field experiments have so far produced mixed evidence that has not consistently indicated specific media violence effects. (pp. 1073–1074)

Somewhere between the two positions presented above, Browne et al. (2005), after reviewing a series of meta-analyses, report consistent effects on arousal, thought, and emotions, culminating in increased likelihood of aggression or fear-induced behavior over the short term in younger children, particularly boys. In older children and teenagers, the evidence for these effects was inconsistent, but there was nevertheless enough of a change in some children to warrant public health concerns. That said, they added that there was only weak correlational evidence to support a link between violent media and crime.

Video Game "Violence"

Over recent years, research on media violence has tended to focus on video game content, particularly depictions and enactments of violence afforded by what Patridge (2013) calls run-of-the-mill first-person shooter games. Anderson et al. (2010), as a result of their meta-analytic review of work published in 2008, claimed to have found that exposure to video games with violent content is a causal risk factor for increased aggressive behavior, cognition, and affect and decreases empathy and prosocial behavior (see also Anderson, 2004). More recently, and again based on a meta-analytic review of published findings (this time between 2009 and 2013), Calvert et al. (2017) likewise reported that exposure to violent video games is associated with increased aggressive behavior, cognitions, and affect, as well as increased desensitization, decreased empathy, and increased physiological arousal. They also report similar effect sizes to prior meta-analyses, which they interpret as indicative of stable results across time. In addition, following a six-month longitudinal study, Greitemeyer and Sagioglou (2017) found that repeated exposure to violent video games predicts everyday sadism (i.e., those who derive personal enjoyment from humiliating or otherwise causing harm to others).

While it would be erroneous to ignore such findings (Greitemeyer & Mügge, 2014), in a similar vein to the exposition of media violence more generally, Ferguson (2007a, b), based on his meta-analytic review of video game violence, warns us to treat many of the results supporting a connection between violent video game content and antisocial behavior with caution, arguing that the measures of aggression used in most studies lack validity and often have effect sizes that are very small (see DeCamp & Ferguson, 2017, for recent findings further supporting the view that exposure to violent video games is not a predictor of youth violence). Ferguson (2007a, b) also suggests that there is a bias in the academic literature in favor of those papers that report statistically significant differences between groups, as opposed to those that do not.[9] Moreover, Markey et al. (2015), following a meta-analytic review of data, including FBI crime statistics and video game sales, report no evidence linking exposure to video game violence to violent crime in the United States (see Cunningham et al., 2016, for similar findings). Finally, and most recently, Przybylski and Weinstein (2019), after conducting an online survey measuring over 1,000 U.K. adolescent gamers and also their respective carers (using a self-report questionnaire), reported "confirmatory evidence that violent video game engagement, on balance, is not associated with observable variability in adolescents' aggressive behavior" (p. 14).

Putting all of this together, at least in terms of research currently undertaken, there is no consensus on what the effects of playing violent video games are (see Ferguson, 2013; but also Bushman et al., 2015; Bushman & Huesmann, 2014; Krahé, 2014, for a rebuttal of Ferguson's 2013 claims, and therefore as a means of reinforcing the argument for a lack of consensus). Therefore, a posteriori, there is no compelling reason (at least where a compelling reason requires a consensus in the empirical findings) to endorse the view that enacting virtual murder (or similar violent killings and/or assaults) is significantly likely to result in harm. Any attempt to posit a direct causal link between video game content and violent (real-world) behavior should therefore be regarded as overly simplistic, largely uncorroborated, and ultimately contentious. Moreover, in accordance with a view expressed by Przybylski and Weinstein (2019), if anything:

> [As a] more nuanced empirical understanding of media effects has emerged, other policy positions, once stridently aligned against gaming and screen time, such as the American Academy of Pediatrics, have softened their prescriptions concerning digital media and psychosocial development. These changes have been reflected in the statutory arena: in 2011, the United States Supreme Court judged that there is insufficient evidence that games cause harm to uphold laws restricting the sale of violent games to minors. (p. 2)

OTHER FICTIONAL SPACES

Research on the negative impact of engaging with fictional immorality (e.g., fictional violence) on affect, cognition, and behavior has tended to focus on film, television, and, more recently, video games. These are not the only fictional spaces in which depictions of fictional immorality are presented, of course, although they have received the most research and scholarly attention. In addition, although by no means an exhaustive selection, the effect on the listener of sexual aggressive, misogynistic, and otherwise violent song lyrics has also been investigated, as has the impact of aggression and violence in novels, particularly those targeting adolescents.

Violent Song Lyrics

In an early study, Barongan and Nagayama (1995) reported that listening to misogynous rap music facilitated sexually aggressive behavior in men toward women. A similar effect was found by Fischer and Greitemeyer (2006), although they did not test for increased sexual aggression, specifically, but aggression measured in terms of the increased ascription of negative attributes by men to women (for example) or the administration of more hot chili sauce to a drink. The latter, in particular, seems to be a less-than-typical measure of aggression, although it has been used elsewhere. Mast and McAndrew (2011), for example, also found that listening to violent lyrics—this time in heavy metal songs—increased the amount of hot sauce participants were prepared to add to water they believed others would have to drink. Further support for a connection between violent song lyrics and aggression was reported by Anderson et al. (2003), based on an increased likelihood to interpret ambiguous words (e.g., alley, animal, stick, bottle) aggressively (see also Brummert Lennings & Warburton, 2011, regarding lyrics and offense). Conversely, prosocial lyrics have been shown to reduce levels of affective, cognitive, and behavioral aggression (Greitemeyer, 2011)

The findings of this relatively small sample of published work appear to support the claim that violent song lyrics have a negative impact on their listener, albeit in a form that one might not typically cite as an example of aggression (although I appreciate that ethical issues no doubt restricted how the researchers were permitted to measure aggressive behavior). That said, Brodsky et al. (2018) found listening to music with violent content increased aggressive and dangerous driving in a driving simulator (which would suggest a higher level of real-world validity). Warburton et al. (2014) offer a word of caution, however, even while reporting increased levels of aggression in adolescents exposed to violent lyrics: that the measured effect found in research on violent lyrics and aggression is usually much smaller than that found when viewing violent television. Likewise, Fatsis (2019) seeks to alleviate U.K. Metropolitan Police concerns over a relatively new form of rap known as drill music, owing to its violent content and perceived association with youth gang violence in London. As Fatsis notes, the lyrical content seeks to depict violent crime (presented in largely fictional form) rather than endorse it: a position no doubt shared by many, if not all, of the composers of the other "violent" song lyrics discussed.

Novels and Aggression

Coyne et al. (2011) performed a content analysis on forty adolescent novels that appeared on the *New York Times* Best Seller List between June and July 2008. They recorded over 7,000 acts of aggression across all the novels. Most were relational or verbal acts of aggression, rather than physical. (Relational aggression is aggression designed to harm a person's relationship, including social standing, within a group.) Coyne et al. concluded that novels are an underresearched medium through which adolescents are exposed to fictional violence. In follow-up research, Coyne et al. (2012) found that an adolescent's level of aggression increased in relation to the type of aggression they had been exposed to in the novel. Those exposed to relational aggression showed increased relational aggression, whereas those exposed to physical aggression exhibited more physical aggression. On a more positive note, although still showing the effects of fiction on the audience, Peterson and Lach (1990) reported that children who read picture books about children successfully combating gender stereotypes were less likely to express stereotypical attitudes about gender.

CONCLUSION

While there has been less research looking into the effects of violent fictional content in media other than film, TV, and video games, what has been carried lends some support to the claim that engaging with depictions of fictional violence and/or other immoral activity through media, such as lyrics, adolescent novels, noncomedic cartoons, and comic books (including manga), increases aggressive behavior and, in some cases, affect and cognition. As with the continuing debate over the effect of exposure to violent content in more researched media, however, such as film, TV, and video games, one may wish to question the way aggression has been measured in some of these studies and therefore how much one can extrapolate from their findings to draw more general conclusions about the effect of a particular medium on aggression.

Where does this leave us on the matter of harm? As noted at the start of this chapter, one may hold that either of the immorality conditions (IC_1 and IC_2) presented in chapter 5 has some merit, *in principle*, while also recognizing that, at present, there is no consistency in the empirical evidence available and, subsequently, no consensus on which if any of these conditions merits support. Given this, one could put the debate over methodology, the operationalization of variables (such as "aggression"), and the interpretation of findings aside and hold, instead, that depictions of fictional immorality, irrespective of the medium through which they are presented, constitute a form of *harmless wrongdoing* (Feinberg, 1988), whereby they are *materially innocent* but *morally non-innocent* (McMahan, 2006)[10] insofar as one might accept that the act of engaging with the fiction is devoid of direct harm (in the absence of evidence to the contrary), making it materially innocent, but nevertheless still consider it to be morally wrong (morally non-innocent). One might come to think of depictions of fictional immorality as contributing, some more than others, to an indirect, perhaps even longer-term, *cultural harm* (McGlynn & Rackley, 2009). Cultural harm, in the context we are discussing, might occur as a result of engendering within a society a *trivializing* attitude toward the immoral acts the fictions depict (e.g., murder, or torture, or rape or child abuse; see chapters 9 and 13 for a return to this point). Borrowing from Oswell (2006), this is because “the ethical intensity of the virtual image lies precisely in its capacity to refer to a scene beyond itself” (p. 258). While the virtual (or other fictional) image or depiction may not be a record of harm, or be shown to contribute directly to new harms (in any agreed manner), nevertheless, how we respond to the depiction, including our *willingness* to engage with it and the *manner* of our engagement, is meaningful, and therefore remains a suitable subject for moral scrutiny.

In the next chapter I examine the morality of *enjoying* depictions of fictional immorality, before moving on to consider, in chapter 8, what, if anything, our *unwillingness* to engage with certain sorts of fictional immorality tells us about the morality of what is depicted.

NOTES

1. Plato’s intellectual criticism of the arts stems from his theory of forms or ideas. For Plato, a perfect circle exists only in the form of an idea. Whenever we recreate a circle, each concrete example is necessarily an imperfect facsimile of the idea of a circle (which is perfect). In *representing* nature (inter alia, on canvass or in stone, through an epic poem and performance of this), art seeks to reproduce the imperfect because nature is imperfect (i.e., nature is a collection of imperfect examples of corresponding ideas: the perfect tree or flower, and so on). As an aside, according to this view, art is seen as a *representation* of nature rather than as an *expression* of the artist’s imagination and creativity. Art is, therefore, twice removed from perfection because it constitutes the reproduction (second copy) of an imperfect (first) copy of the perfect idea. Consequently, engaging with art twice removes us from knowledge and understanding of this perfect idea, which, instead, we must achieve through intellectual pursuit in the form of philosophical engagement.

2. https://www.dailymail.co.uk/news/article-2881705/Descendants-notorious-19th-Century-writer-Marquis-Sade-reclaim-title-200-years-disowning-novelist.html (assessed 7/12/19). Also, in 2016, Penguin Classics published *120 Days of Sodom*.

3. The current standard by which obscenity is measured in the United States—often referred to as the *community standard test*—is taken from *Miller v. California* (413 U.S. 15) in 1973 and comprises the following three components or tests, which, if met, classify the object/event as legally obscene: (i) whether “the average person, applying contemporary community standards” would find that the work, taken as a whole, appeals to the prurient interest; (ii) whether the work depicts or describes, in a patently offensive way, sexual conduct specifically defined by the applicable state

law; and (iii) whether the work, taken as a whole, lacks serious literary, artistic, political, or scientific value (see Feinberg, 1988, for a detailed discussion).

4. F3d 572 (7th Cir 2001).

5. What is judged to be depraved and corrupting must also be weighed against expert opinion regarding the extent to which the material "is justified as being for the public good on the ground that it is in the interests of science, literature, art or learning, or of other objects of general concern" (Obscene Publications Act, Section 4:1).

6. In fact, Stone (2007) cites the case of *Paris Adult Theatre v Staton*, in which the court ruled that obscenity may be regulated in part because it is harmful.

7. See also Parents Television Council's Special Report (2013) Media violence: An examination of violence, graphics violence, and gun violence in the media. http://w2.parentstv.org/main/Research/Studies/CableViolence/vstudy_dec2013.pdf (accessed June 17, 2019).

8. It is worth noting that reference to cathartic purging in fact relates to a conception of catharsis different to that originally used by Aristotle. In this case the idea of catharsis as a *purging* of emotion is something that can be traced to Freud. Recall that, for Aristotle, catharsis describes the purification of emotion—fear and pity—not its purging (see Gentile, 2013, for a detailed discussion).

9. By way of additional dissenting voices and further critical discussion on Anderson et al.'s (2010) conclusion, see Bushman et al. (2010), Ferguson and Kilburn (2010), and Huesmann (2010). See also Bensley and Van Eenwyk (2001) and Ferguson (2011).

10. McMahan used these terms originally in relation to his critique of just war theory.

Chapter 7

Enjoying Fictional Immorality

INTRODUCTION

In the previous two chapters, I considered the question of harm: its relationship to engaging with depictions of fictional immorality and the extent to which research supports the claim that engaging with fictional-x_i has a deleterious effect on oneself and others. While the consequences of engaging with fictional-x_i warrant moral appraisal for all the reasons discussed, at the same time, one needs to be cautious when scrutinizing effect alone, especially if one's scrutiny is based solely on the fact that someone is *willing* to engage with depictions of fictional immorality (Bartel, 2015). In addition, the *reason* the person is willing to engage with (inter alia) fictional murder or assault, torture, or rape needs to form part of one's moral appraisal.

In this chapter I examine whether *enjoyment*, as a reason for engaging with depictions of fictional immorality, is an immoral thing to do. (As an aside: I focus on enjoyment while recognizing that there are other motivations/reasons for willingly engaging with fictional immorality.) My response is somewhat nuanced. I argue that the morality of one's enjoyment is dependent on the type of enjoyment one is motivated to procure but, also, that this type of enjoyment shapes the moral nature of the fictional space one's enjoyment creates. When one's enjoyment of fictional immorality is directed at the depiction itself, it is my contention that the immorality of what is being depicted is not *sufficient* for one's enjoyment to be considered immoral, including the fictional space one's enjoyment creates. Where the fiction is used to enjoy vicariously what the depiction is a depiction of, however (e.g., actual torture, rape, or murder), then one's enjoyment, *in this sense*, along with the fictional state created in order to procure this type of enjoyment, ought to be deemed immoral. This is because one should not enjoy (albeit in a manner yet to be discussed) that which is immoral, nor, subsequently, should a fictional space be created for the procurement of this type of enjoyment.

ENJOYING FICTIONAL IMMORALITY

I take the following proposition to be true with qualification:

P_0: *Enjoying* ϕ is immoral *if* ϕ is immoral.[1]

I say "with qualification" because the truth of P_0 is dependent on the type of enjoyment referred to within the proposition. For now, let us align this with Ostritsch's (2017) *strong* sense, which he equates to something being *cheerful* or *fun*, even *thrilling*. In accordance with P_0 and Ostritsch's definition, then, it would be immoral to be thrilled by the abduction of an innocent or to be cheered by news of their murder, or to be amused by the later revelation that they had been tortured to death. That said, as we progress, it is my intention to refine Ostritsch's definition in order to differentiate between two types of enjoyment, each compatible with Ostritsch's strong sense but which affects the truth of P_0 (and related propositions) differently. For now, however, it will suffice to utilize Ostritsch's strong sense of enjoyment until otherwise stated. Thus, enjoying watching a documentary on the war crimes of World War II Japanese experimental biological and chemical warfare unit, 731, because one is amused by the stories of torture and suffering, as with the previous example of the abduction, torture, and murder of an innocent, is to enjoy immorality in a manner consistent with the sentiment of P_0: that is, in an immoral way.

Consider, then, proposition P_1:

- P_1: Where ϕ is immoral, it does not follow (logically, conceptually, or causally) that "other than ϕ" is immoral
- P_1 is unremarkable. A particular instance of P_1 is p_1.
- p_1: Where x is immoral (x_i), it does not follow (logically, conceptually, or causally) that fictional-x_i is immoral.

Recall that fictional-x_i *depicts* an immoral action (x_i) that is not directed toward, nor does it involve in any significant way, the depiction of an actual object, animal, person, or event, or otherwise allude to any of these. The murder of the fictitious character R. K. Maroon in the 1988 film *"Who Framed Roger Rabbit?"* is an example of fictional-x_i, whereas the murder of real-life murder suspect Lee Harvey Oswald as depicted in Oliver's Stone's 1991 film, *JFK*, is not.

Like P_1, proposition p_1 is hardly remarkable or contentious. The reason for its inclusion will, however, become apparent as we progress. In support of p_1, recall from chapter 3, Di Muzio's (2006) assertion that "depictions of violence do not per se belong in the category of the morally objectionable only because many instances of real violence do" (p. 280). Thus, Di Muzio continues, "one would want to resist the thesis that it is wrong to read and enjoy Homer's *Iliad* because it contains violence, gore and death" (*ibid.*). In light of p_1, consider the following proposition:

> p_2: *Enjoying* fictional-x_i is immoral because that which fictional-x_i depicts (namely, x_i) is immoral.

According to p_2, it is immoral to enjoy enacting the murder of a character in a video game, or to enjoy watching the torture of the police officer in *Reservoir Dogs*, or to find a racist joke about (inter alia) a token and abstracted black or Irish or Jewish person funny, or even to enjoy watching the exploits of a dandy highwayman or a ruthless contract killer. You may find some of these examples easier to align with p_2 than others. Perhaps, intuitively, the strongest support for p_2 occurs on those occasions when S is said to enjoy depictions of fictional rape or pedophilia or torture: say, while playing a video game containing such enactments (e.g., *RapeLay* and *Enzai: Falsely Accused*) or when watching a film like Sam Peckinpah's *Straw Dogs* or the unlicensed *A Serbian Film*. Under such circumstances, one might declare, in accordance with p_2: "Enjoying depictions of fictional rape is immoral precisely *because* rape is immoral." *Mutatis mutandis*, the same can be said about fictional pedophilia or torture. Despite acknowledging the force of such an intuition, for p_2 to have any appeal beyond an intuitive one, its truth must be reconciled with the truth of p_1: the somewhat unremarkable and uncontentious proposition that the immorality of x_i does not entail the immorality of fictional-x_i.

To understand why, consider the following conjecture: If we accept p_1, then we are accepting that the moral status of x_i is not sufficient to determine the congruent moral status of fictional-x_i (a not unreasonable position to adopt). If, however, we accept p_2 then we hold that the moral status of x_i is sufficient to determine the congruent moral status of one's enjoyment of fictional-x_i. In other words, where one accepts the truth of p_1 *and* p_2, then one accepts that the immorality of x_i is not sufficient for fictional-x_i to be immoral but is sufficient for one's *enjoyment* of fictional-x_i to be immoral. In effect, one is declaring that it is immoral to enjoy fictional-x_i even when fictional-x_i is not immoral, which is essentially saying that there are times when it is immoral to enjoy that which is not immoral.

Given the unpalatability of such an outcome, does it mean, at the very least, that either p_1 or p_2 is false? In chapter 3, I argued in favor of p_1. I also stated earlier that, intuitively, some examples of enjoying fictional-x_i align more readily than others with p_2. In order to accept both p_1 and p_2, then, we must:

1. Allow that p_1 is depiction-dependent and therefore false at least some of the time, depending on what is being depicted (e.g., where what is depicted is, inter alia, torture or rape or pedophilia).
2. Allow that p_2 is likewise depiction-dependent and therefore false at least some of the time (say, in cases of enjoying depictions of fictional theft or murder).
3. Claim that, as presented, p_1 and p_2 are susceptible to the charge of equivocation and thus seek to reformulate each proposition in a manner that (a) dissolves the equivocation, and (b) allows p_1 and p_2 to be both true and not depiction-dependent.

In accordance with my argument in chapter 3, I reject option 1 unconditionally and, with it, the truth of the claim that the immorality of a depiction of fictional immorality (*qua* pure fictional immorality) is depiction-dependent, and therefore that some depictions of fictional immorality are inherently immoral. Regarding option 2, I have experienced the thrill of watching a film featuring a charismatic bank robber or cat burglar take what is not theirs in some high-tech, adrenaline-fuelled manner, without considering my enjoyment to be immoral, and found myself amused by the antics of a contract killer—say, in a film like *Grosse Pointe Blank*—again, without experiencing any sense of impropriety. Consequently, the depiction-dependency mentioned in option 2 does *seem* to carry some weight: to have some intuitive appeal. But can my intuition that what I am doing is not immoral be secured to a more solid theoretical base: one that offers a means of distinguishing between the enjoyment of depictions of fictional immorality that is immoral and enjoyment that is moral, or at least not immoral? Alternatively, might a closer examination of my intuition reveal it to be mistaken, insofar as I am not entirely without fault, morally, to cheer on and be cheered by the exploits of the bank robber or be amused by the assassin, given that what each does is immoral when carried out for real? If the answer is "yes," then is the difference between enjoying the exploits of the thief, compared to the torturer, one of degree rather than kind? In other words, should the immorality of my enjoyment correlate with the immorality of the act when done for real, meaning that my enjoyment of fictional theft should be less harshly judged than my enjoyment of fictional torture or rape or pedophilia because, typically, the latter acts are considered to be a greater immorality than theft? Even *if* this were the argument, each action is still nominally immoral and so fails to show that p_2 is ever *false*.

Perhaps, instead, we could claim that all enjoyment of fictional-x_i is immoral *in principle* but maintain that only the enjoyment of fictional immorality *of a certain kind* is worthy of note and so in need of overt condemnation. To illustrate, perhaps my enjoyment of the antics of hit man Martin Blank (in the film, *Grosse Pointe Blank*) should be treated with less moral outrage than my enjoyment of the exploits of John Kramer (aka Jigsaw) in the *Saw* franchise. After all, different narratives and genres of fiction provide different context-framing for the meaning of violence (Symonds, 2008).

In contrast to both of these options, the remainder of the chapter is dedicated to defending option 3, but in a manner that allows some light to shine on the seeming depiction-dependency evident in 2. In what is to come, I examine the relationship between x_i and fictional-x_i in order to identify and resolve the equivocation that I believe exists between p_1 and p_2, at least in their current formulation. In order for this to be done, and for p_1 and p_2 to be accepted without fear of incoherence, the following must occur:

- Ostritsch's (2017) strong sense of enjoyment must be refined.
- Fictional-x_i must be adapted in accordance with Walton's (1990) work world/game world distinction as a means of producing ontologically distinct forms of fiction.

EXAMINING THE RELATIONSHIP BETWEEN X_I AND FICTIONAL-X_I

Fictional-x_i's ability to depict immorality is dependent on the moral status of x. If x is immoral (x_i), then necessarily fictional-x_i *depicts* immorality (I say this in the absence of any comment about how well—either in terms of aesthetics or in terms of representational realism—fictional-x_i depicts x_i). Given this, let us consider the effects of this relationship on:

(a) The autonomy of S's *understanding* of fictional-x_i. That is, the extent to which S's understanding of what she is doing when performing fictional-x_i is independent of her understanding of x_i.
(b) The autonomy of S's *enjoyment* of fictional-x_i. That is, the extent to which S's enjoyment of fictional-x_i is independent of her enjoyment of x_i.

Ultimately, this is done to qualify:

(c) The dependency relation between the morality of one's enjoyment of fictional-x_i and the morality of x_i.

Understanding

Suppose S enjoys playing video games that afford the opportunity to enact murder. To what extent is S's predilection for these games dependent on actual murder? In order for S to understand that what she is enjoying is an enactment of *murder*—when she finds herself enjoying *this* particular enactment—then, necessarily, she must have some understanding of what murder is—what it might look like—and therefore what is involved in performing the action. S must also understand that what is being depicted is intended to depict murder. More formally:

- S understands that fictional-x_i is an enactment of x_i *if* and *only if* S understands that the actualization of what fictional-x_i depicts would be a token of x_i.

To be clear, by "understands" I mean more than S is simply able to *recognize* that fictional-x_i depicts something called "murder," despite failing to be adequately cognizant of what murder is. Of course, given the range of ways in which murder can occur, it is unlikely that S would be able to recognize more than a few examples of murder without having an adequate grasp of what murder entails. Moreover, I also accept that S may reimagine a virtual enactment that is not murder as an enactment of murder. Ali (2015) provides a useful example of this when describing a hypothetical gamer who enacts the killing of a character in self-defense as an act of murder designed to look like self-defense. I will have more to say on this possibility when discussing Walton's work world/game world distinction (see the section "Refining Fictional- x_i: Distinguishing between Fictional Worlds") and how each constitutes its own fictional world and therefore, where applicable, its own fictional-x_i. With these caveats in place, then, I consider the proposition above to be an adequate description of the necessary and sufficient conditions for understanding a depiction of fictional-x_i.

Enjoyment

Where x_i constitutes a token act of murder, while it may be the case that S's understanding of fictional-x_i—namely, that it is an enactment of *murder*—is dependent on S's understanding of what constitutes (actual) murder, the same dependency relation is absent in the case of one's *enjoyment* of virtual murder. In other words, while one could not have an understanding of what virtual murder (m_v) is meant to be depicting without an understanding of what fictional-x_i is a depiction *of* (namely, actual murder: m_a), the same dependency relation does not hold in the case of enjoying m_v. While one's understanding of what m_v is depicting necessitates understanding m_a, enjoying m_v does not necessitate enjoying m_a (or vice versa). In fact, not only is one's enjoyment of m_v independent of enjoying m_a, it is independent of one's understanding of m_a. After all, when playing a video game (for example), it is not inconceivable that one might enjoy a particular enactment on a given occasion—owing, say, to its elaborate graphics and the spectacular nature of the avatar moves—without understanding that what one has just enjoyed enacting (fictional-x_i) is a brutal murder. But, equally, it is conceivable that, in addition to the enjoyment elicited by elaborate graphics and spectacular avatar moves, one's enjoyment of fictional-x_i is heightened precisely because one understands that it is an enactment of murder. More formally, one might surmise the following:

(1) S has the capacity to enjoy (develop a predilection for enjoying) m_v in the absence of any understanding that what she enacts (in the case of m_v) is an enactment of murder.

Of course, in the case of (1), S would not be enjoying m_v *as* m_v (i.e., as a depiction of m_a), given that she does not understand (at least, on this occasion) that m_a is what fictional-x_i is depicting. However:

(2) S's enjoyment of m_v may well be enhanced by the fact that she understands that what fictional-x_i depicts is an act of *murder*.

It is important to note that (2) asserts only that one's enjoyment *could* be enhanced in the way described; it does not rule out the possibility that S may enjoy m_v simply for the reasons mentioned earlier (i.e., elaborate graphics and spectacular avatar moves), even while fully cognizant of the fact that she is enacting murder. In addition, and importantly, (1) and (2) are compatible with the following:

(3) Irrespective of whether S understands that what she is doing is enacting murder, S has the capacity to enjoy (to develop a predilection for enjoying) m_v in the absence of any enjoyment (including delighting in the idea) of m_a.

For S to enjoy m_v, in accordance with (1), S need not understand that m_v is an enactment of *murder*. Where S does not approve of m_a, and subsequently (let us allow) does not enjoy m_a (in accordance with (3)) but, nevertheless, enjoys m_v (in accordance with either (1) or (2)), then S must enjoy m_v for a reason other than enjoying m_a, because, as we have seen, S lacks such enjoyment. Necessarily, S must enjoy m_v for some other reason.

One such reason is that S enjoys m_v for its own sake. In the case of (1), this simply means in virtue of being some kind of *enactment* (perhaps with the aforementioned elaborate graphics and so on). Given that (1) does not require an understanding of the enactment *as* an enactment of murder, I shall dispense with this possibility, as I consider it to be of limited interest to my argument.[2] What is of interest, however, is the idea espoused by (2): that S's enjoyment of virtual murder may be *enhanced* by the fact that she understands that what she is enacting is murder—therefore, not for purely aesthetic reasons. On such an occasion, S's enjoyment of fictional-x_i, *for its own sake*, amounts to not only the enjoyment of m_v as an *enactment* (a not unimportant fact, as I intend to show) but also as an enactment of *murder*: something she understands to be immoral—prohibited, a taboo—when done for real.

To illustrate the relationship alluded to here, consider the weaker or minimal sense of enjoyment (enjoy_w) used by Ostritsch (in contrast to the stronger sense introduced earlier) whereby one is captured or fascinated by something immoral without one's enjoyment being likewise considered immoral. One might, for example, enjoy_w learning about mediaeval torture techniques because one is interested in the psychology of interrogation or enjoy_w reading *Mein Kampf* because one wishes to understand the appeal of Nazism or Hitler's use of rhetoric; but, in neither case should one's enjoyment_w of the immoral be taken to signify one's endorsement of it (recall discussion on endorsement in chapter 4). For this reason, the criminal psychologist, who is fascinated by the "mind" of the infamous child murderer Ian Brady, should be able to enjoy_w learning about motivations for child torture and murder without fear of rebuke. Importantly, though, where one enjoys_w studying the writings/speeches of Hitler or the mind of a serial killer, arguably, in each case, the target of one's enjoyment_w is not *this* particular immorality (whatever it happens to be) but, rather, the act of *reading* or *studying* or otherwise *learning* about this particular immorality. Consequently, one does not enjoy_w the immorality per se. I am willing to concede that there is some truth to this claim. Nevertheless, we must not lose sight of the fact that the enjoyment_w one experiences in such cases is unlikely to be derived solely from the act of studying but, rather, from the combination of engaging with *this* particular immoral activity in *this* particular way (i.e., *understanding* the mind of a *serial killer* as opposed to, say, a butcher, baker, or candlestick maker). One's enjoyment_w may also stem, in part or whole, from the originality of the point of view and/or sentiment depicted within the fiction: the viciousness found in Sade's *120 Days of Sodom*, for example, may elicit its own intellectual, including aesthetic, fascination (Gaut, 2007). The same applies to one's enjoyment in the stronger sense, I contend: that one enjoys *this* particular immoral activity in *this* particular way.

In the case of video game violence, S's enjoyment is heightened not just because she is engaged in some form of enactment or simulation (even though this can have its own sort of appeal, as (1) attests) but because she understands that what she is doing is virtually enacting *murder* (Nys, 2010). Fictional-x_i is therefore understood to be a means of depicting something known to be *taboo*. It is worth stressing, however, that one's enjoyment of fictional-x_i is heightened not simply because one understands that what one is doing (in this case) is enacting *murder* but because, underlying this, one understands that what one is doing is *enacting* murder (i.e., *not* doing it for real). In accordance with (2), then, enjoying fictional-x_i, for its own sake, means deriving enjoyment from what one does when what one does is understood to be an *enactment* (not the actualization) of a *taboo*.

Where one enjoys the *enactment* of murder because it is an enactment of *murder*, while also not enjoying actual murder (viewing it with distain, let us say), the respective states of enjoyment and distain are logically and ontologically independent because the objects of one's enjoyment and distain are logically and ontologically independent (it is hardly contentious to assert that fictional-x_i and x_i are not the same). What remains to be discussed, however, is how far these two objects, and the emotional responses they elicit, are psychologically (dis)connected.

SIMULATING SADOMASOCHISM

In light of the discussion thus far, consider the claims made by Hopkins (1994) about those who enjoy participating in acts of sadomasochism (SM):

> In the case of SM . . . it should not be assumed that SM participants actually find pleasure in the torture of slaves, nor in the cries of a rape victim, nor in the humiliation of women, nor in the relentless assault of an attacker. . . . At the same time, however, it is possible to desire the *simulation* of those events, to lust after the context of a negotiated and consensual "submission" or "domination". . . . This does not mean that simulation is the closest the SM practitioner can get to her real desires. This does not mean that the simulation of rape is a legal stand-in for the real thing. Neither should it be taken for granted that the participants get their pleasure by getting so far into the fantasy that they feel like it is the real thing. Rather, the sadomasochist can *desire the simulation itself*, not as an inferior copy of the real thing, not as a copy of anything at all, but as simulation qua simulation. (p. 125; emphasis in original)

Hopkins' claim about the SM practitioner's desire for simulation *qua* simulation accords with (2) insofar as the practitioner desires not only simulation *qua* simulation but, importantly, simulation *qua* the simulation *of* certain *immoral* practices (e.g., slavery and/or rape), *understood* to be the simulation of certain immoral practices. (Note: I intend to use the term "simulation" as a synonym of "enactment," something I consider to be compatible with my liberal use of the term "depiction"). To acknowledge this fact is to acknowledge that enjoying a simulation in this way cannot be divorced from the practitioner's *understanding* of what the simulation depicts, but this is not the same as saying that one's *enjoyment* of the simulation necessitates one's enjoyment of what the simulation depicts (although see below for a contrary view). This last assertion accords with Hopkins' further claim that the simulation is not desired because it acts as a substitute for the real thing: that is, as a substitute for what the practitioner desires more than the simulation itself but which is prohibited (e.g., to enslave and/or rape). Such a position is, of course, suggestive of (3) insofar as the SM practitioner is able to enjoy the simulation in the absence of any (anticipated) enjoyment of the real thing, including delighting in the idea of the real thing.

Corvino (2002), in contrast, has this to say: "The simulation is not the object of arousal; rather, it is the *vehicle* for the object of arousal" (p. 215; emphasis added). What elicits one's enjoyment, according to Corvino, "is not depiction, but rather what is being depicted" (*ibid.*). The simulation/enactment/depiction acts as the vehicle or means by which one derives enjoyment from x_i in the absence of x_i. Fictional-x_i is not enjoyed for its own sake; rather, it functions as a substitute for what one does desire but cannot attain (see also Kjeldgaard-Christiansen, 2020, who makes a similar distinction in relation to our intuition about enacting virtual pedophilia compared to virtual murder). Of course, there is no reason why both authors cannot be correct, depending on which particular SM practitioner one has in mind. S_1 may enjoy engaging in sadomasochistic practices for the reason espoused by Hopkins. S_2, however, may do so for a reason approaching Corvino's explanation. Assuming both of these authors have identified a valid motivation, identifying which desire applies to which practitioner is ultimately an empirical matter.

DIFFERENTIATING BETWEEN FORMS OF ENJOYMENT

In previous work, I have posited a motivation for enacting taboos that is compatible with a number of components of the discussion so far: Ostritsch's strong sense of enjoyment, Hopkins' explanation, and point (2) within this chapter. In Young (2013), I call the motivation, $M_{(enjoyment)}$. Where S enacts fictional-x_i in accordance with $M_{(enjoyment)}$, S enjoys fictional-x_i because it is an enactment of a *taboo*. As we have seen, point (2) is compatible with point (3), thereby making $M_{(enjoyment)}$ compatible with point (3). Fictional-x_i is therefore something S can be motivated to do in the absence of deriving enjoyment from x_i. But what of the person who employs fictional-x_i in a manner consistent with Corvino's explanation? In Young (2013), I refer to this motivation as $M_{(substitution).}$ Here, S is motivated to engage with fictional-x_i as a substitute for what they actually desire and are motivated to enjoy: namely, x_i.[3] To illustrate, suppose S enjoys playing the Japanese video game *RapeLay*.[4] According to $M_{(substitution)}$ and also Corvino, the *enactment* of rape is not necessarily enjoyed for its own sake but because of what it represents (i.e., actual rape). The simulation acts as a substitute for the act of rape. Typically, a rapist who enjoys rape enjoys something that *happens to be taboo*; they do not enjoy it simply because it is taboo: for *that* reason. Where the rapist enjoys enacting rape, it is because (one might reasonably conjecture) it depicts that which he already enjoys (or, in the case of the would-be rapist, delights in the idea of). The rapist therefore attains vicarious enjoyment through a game like *RapeLay*. In contrast, where one derives enjoyment from the enactment, *for its own sake*, then one's enjoyment would not be vicarious but direct. The person who enjoys *enacting* rape, for its own sake, according to $M_{(enjoyment)}$ and Hopkins, enjoys enacting that which is taboo because it is a *taboo*. *Simulating* its taboo-ness *is* the attraction, which may be all the more enjoyable if the simulation also has a certain aesthetic appeal.

Ostritsch's strong sense of enjoyment captures, without differentiation, what I intend to call enjoyment *qua* simulation and enjoyment *qua* substitution—each of which amounts to a unique motivation with its own moral implications. Ostritsch's definition is therefore too broad and needs refining.

In accordance with my distinction between player motivations, enjoyment *qua* simulation (henceforth "enjoyment$_{(sim)}$") denotes enjoyment elicited by the simulation itself. In the case of murder (for example), if one enjoys$_{(sim)}$ enacting murder, then what one enjoys—finds cheerful, fun, or thrilling—is enacting a taboo as an *enactment* of a *taboo*. In contrast, enjoyment *qua* substitution (henceforth "enjoyment$_{(sub)}$") denotes one's enjoyment of what the depiction is a depiction of. Someone who enjoys$_{(sub)}$ enacting murder does not necessarily enjoy—find cheerful, fun, thrilling—*enacting* murder as an end in itself but, rather, enjoys this activity as a means to an end or as a vehicle (to borrow Corvino's term) through which they can vicariously satisfy what they do desire and/or delight in the idea of (i.e., actual murder). To say that S finds depictions of fictional rape (as found in, say, the film *Straw Dogs* or the video game *RapeLay*) enjoyable$_{(sub)}$ is to say that S is cheered by, or finds funny, or is thrilled by the depiction not for its own sake—or not necessarily for this reason—but *necessarily* because, for S, it provides a means of procuring vicarious satisfaction. The reason S finds enacting rape in the video game

RapeLay thrilling, then, is because he finds what the enactment is an enactment of (or at least the idea of what it is of) thrilling. This is in contrast to the person who is amused, cheered, or thrilled by the *enactment* itself. That is, who is amused, cheered, or thrilled by the opportunity to *enact* (or otherwise engage with) a *taboo*.

One might object, however, to such a description by arguing that someone who enjoys the depiction of rape in *Straw Dogs*, even as a *depiction* only (i.e., enjoys$_{\text{(sims)}}$), is missing the point of the scene and is reacting in an inappropriate way. This would not be the case, or would seem to be far less the case, if one were to enjoy$_{\text{(sim)}}$ *RapeLay*. I will have more to say on authentic and inauthentic engagement when discussing Walton's work world/game world distinction. As a precursor to this, in the case of *Straw Dogs*, even if the director explicitly states that he was not seeking to elicit enjoyment$_{\text{(sim)}}$ from the audience (during the rape scene), where S reacts in just such a way, is the misalignment—the incongruence between intent on the part of the director and S's reaction—immoral? I will address this question below.

REFINING FICTIONAL-X_I: DISTINGUISHING BETWEEN FICTIONAL WORLDS

With the enjoy$_{\text{(sim)}}$/enjoy$_{\text{(sub)}}$ distinction in place, I present two variations on p_2. It is my contention that the first (p_2*) is false, whereas the second (p_2**) is true insofar as it amounts to a rudimentary version of what, with further refinement (in order to avoid still more equivocation), is a defensible position.

p_2*: *Enjoying*$_{\text{(sim)}}$ fictional-x_i is immoral because that which fictional-x_i depicts (namely, x_i) is immoral (false).

p_2**: *Enjoying*$_{\text{(sub)}}$ fictional-x_i is immoral because that which fictional-x_i depicts (namely, x_i) is immoral (true with refinement).

The reason p_2** is true (even though it needs refinement) is because, when enjoying$_{\text{(sub)}}$ fictional-x_i, the fiction the *agent creates* in order to procure enjoyment$_{\text{(sub)}}$ is necessarily immoral. Such an outcome is therefore compatible with P_0 (recall: *Enjoying* ϕ is immoral if ϕ is immoral). In order to defend this claim, however, I must distinguish between what Walton (1990) refers to as *work worlds* and *game worlds*.

A work world, broadly construed, is the authorized fiction: A long time ago, in a galaxy far, far away, there existed a Rebel Alliance and a Jedi knight named Luke Skywalker (and so on). Work worlds are said to consist of fictional truths conducive to all *authorized* game worlds. Game worlds therefore consist of make-believe truths that stem from the fictional content of the work world. One's game world may be authorized—that is, closely aligned with the fictional truths contained within the work world—or unauthorized, amounting to a subversion of this fictional truth (see Ali, 2015, for further discussion). Thus, one may view the film *Star Wars* or even play the video game while fictively imagining that the Rebel Alliance consists of freedom fighters opposing an evil empire (authorized game world) or, instead, fictively imagine that the Rebel Alliance is, in fact, just a bunch of fanatical terrorists out to pervert the natural order of the galaxy (unauthorized game world).

Where fictional-x_i *qua* the work world is not immoral, even if what the work world depicts is (e.g., murder)—a position that accords with p_1—then, when one engages with the work world, in order to elicit enjoyment$_{(sub)}$ from it, the fictional world one creates as a result of this engagement—one's own unique *game world*—is immoral. It is immoral not because it amounts to an unauthorized version of the work world—it may or may not do this[5]—but, rather, because the game world's creation—its coming into existence—is the direct result of the agent's *motivation* to use the work world as a vehicle to satisfy, vicariously, their desire to engage with *x* or at least delight in the idea of this. A consequence of one's motivation to treat the work world in this way, then (i.e., as a vehicle), is the act of enjoying$_{(sub)}$ an immoral game world. The latter outcome is therefore a product of the former motivation. In contrast, and returning to the question I posed earlier, if one were to enjoy$_{(sim)}$ the rape scene in *Straw Dogs*, and even if (as is likely the case) such enjoyment is not a reaction intended by the director, then this incongruence, in and of itself, does not make one's reaction immoral; it simply means it is out of kilter with the reaction the director sought to elicit. Saying this does not absolve S of wrongdoing altogether, however—there is still the matter of moral insensitivity in the form of poor taste to consider—but it does counter any claim that enjoying$_{(sim),}$ *qua* an unauthorized game world, the rape scene in *Straw Dogs* is immoral.

The distinction between work worlds and game worlds makes p_2** compatible with p_1, as can be seen here:

> p_2**$_{\text{REFINED}}$: S's enjoyment$_{(sub)}$ of fictional-x_i (*qua* their own unique game world) is immoral because that which fictional-x_i depicts (namely, x_i) is immoral, even where fictional-x_i (*qua* work world) is not immoral.

By refining the type of enjoyment it is morally problematic to enjoy (i.e., enjoy$_{(sub)}$), and stipulating the type of fictional state this enjoyment applies to (i.e., the agent's game world), not only is a sufficient condition for the immorality of one's enjoyment revealed but also the reason for its immorality. Where fictional-x_i *qua* work world depicts immorality in virtue of the fact that one enjoys$_{\text{(sub)}}$ this immorality, then the game world one produces in order to procure this enjoyment$_{\text{(sub)}}$ is immoral *because* it is a product of vicarious wish fulfillment that is itself immoral. A more causally direct way of describing the relationship is to state that enjoying$_{\text{(sub)}}$ one's game world is immoral on this occasion *because* one's game world is immoral on this occasion, and so one is enjoying that which is immoral (thereby aligning $p_2{}^{**}{}_{\text{REFINED}}$ with P_0). Moreover, one can say that this relationship is necessarily the case whenever fictional-x_i *qua* work world depicts the immoral, *and* one positions fictional-x_i *qua* work world as a vehicle for procuring enjoyment$_{\text{(sub)}}$ of x_i. A *necessary* consequence of the process of procurement described *is* the creation of an immoral game world: a world created for the purpose of delighting in the idea of (a specific type of) real-world immorality.

It is important to note that I make no claim that x has to be immoral for one's enjoyment$_{\text{(sub)}}$ to be immoral. Here, I am concerned with the immorality of x (*qua* x_i) as a sufficient condition, not a necessary one. Where x is not immoral, where the work world is not (intended to be) immoral, one's enjoyment$_{\text{(sub)}}$ may nevertheless be immoral in accordance with P_0 (again, recall: *Enjoying* ϕ is immoral if ϕ is immoral). For an occasion on which one's game world would be immoral, even where x or the fiction *qua* work world is not, imagine that S enjoys$_{\text{(sub)}}$ watching the musicals *Oliver* and *Annie* (each of which features a number of minors in lead and supporting roles) as vehicles to delight in the *idea* of child solicitation and molestation and, in doing so, creates an unauthorized fictional world *qua* game world to procure such enjoyment$_{\text{(sub)}}$.

Leaving this last point aside, the seeming depiction-dependency of p_2 (as noted in section "Enjoying Fictional Immorality")—that it is only true in the context of certain fictions such as rape or pedophilia or torture—may be explained in terms of the tendency to ascribe to those who enjoy depictions of rape or pedophilia or torture a motivation akin to enjoyment$_{\text{(sub)}}$, whereas those who enjoy murder (for example) a motivation akin to enjoyment$_{\text{(sim)}}$ (see Ali, 2015, for a related discussion; also, Ramirez, 2020). Whether such an ascription bears any resemblance to fact is, however, an empirical matter and, at best, a contingent fact. As such, it does not negate the possibility that others could enjoy$_{\text{(sub)}}$ murder or theft, or enjoy$_{\text{(sim)}}$ rape or pedophilia or torture, thereby making x_i *sufficient* for the immorality of one's enjoyment in the former case but not the latter.

To explain this last remark: S's enjoyment$_{\text{(sim)}}$ of fictional-x_i (*qua* their own unique game world) is *not* immoral if x is immoral, even in cases of rape, pedophilia, or torture. This is because when procuring enjoyment$_{\text{(sim)}}$, the game world S creates in order to do this is not *necessarily* immoral. It is not immoral simply because x_i is. Differentiating between enjoyment$_{\text{(sim)}}$ and enjoyment$_{\text{(sub)}}$ is therefore unable to make true the depiction-dependency of

p_2 (as was suggested earlier when comparing cases of depictions of fictional rape with fictional murder): for even where x_i remains constant (e.g., murder), there will be occasions when fictional-x_i is enjoyed in a moral (or at least not immoral) or an immoral way depending on the type of enjoyment one is motivated to procure (i.e., enjoy$_{(sim)}$ or enjoy$_{(sub)}$) (see Smuts, 2016, for a discussion on the morality of the fictional *experience*). Importantly, though, this does not negate the possibility that enjoying$_{(sim)}$ depictions of fictional child sexual assault or rape or torture is immoral for some other reason, or at the very least a candidate for poor taste (see Patridge, 2013; Young, 2017, 2019; see, also, chapter 9 for further discussion on this point).

The fact that enjoying$_{(sim)}$ fictional depictions of certain immoral actions may be perceived by many as more morally troubling than others (e.g., rape compared to murder), and perhaps, at the very least, as a candidate for poor taste, aligns with the seeming depiction-dependency of p_2 (noted earlier). Of course, it may be that there are other factors, yet to be considered that can account for the depiction-dependency intuition. That said, what I am attempting to show, here, is that p_2 is false, except when formulated as $p_2{}^{**}{}_{\text{REFINED}}$. To further support this conclusion, I will finish by applying $p_2{}^{**}{}_{\text{REFINED}}$ to a less discussed form of depiction: singing.

SINGING ALONG TO DEPICTIONS OF FICTIONAL IMMORALITY

Smuts (2013) discusses the morality of singing along to certain songs. In particular, he draws our attention to a song by the rap group *Geto Boys*, entitled *Mind of a Lunatic*. The song depicts (fictional) rape, murder, and necrophilia. Smuts argues that when one sings along with a song like *Mind of a Lunatic*, one tends to visualize acting out the content. (Note: While the extent to which one "tends" to do this is yet to be substantiated, empirically, it nevertheless remains something that it is possible to do.) Consequently, the song provides an opportunity for the listener to enjoy (from a first-person perspective) fictively imagining doing "evil" (Smuts' term): that is, performing grossly immoral acts.

In the case of *Mind of a Lunatic*, the work world involves the fictional truth that one is able to engage in various immoral and brutal acts (as described). By singing along, as with the previous examples of film and video game depictions, one constructs, through one's fictive imagination, a game world. Importantly, what one is enjoying when one sings along is contained within the fictional world of one's imagination now avowed in virtue of one's sing-along.

Consider, then, S_1 who enjoys the immorality afforded by the song's work world. S_1's game world consists of make-believe truths about "his" interactions with fictional characters (he make-believes *qua* fictively imagines that he is a rapist, and so on). What he enjoys is engaging in these *make-believe* truths about *fictional* characters. The objects of his enjoyment are found within his game world, and these objects are depictions of fiction. He enjoys *playing the part* of a rapist, murderer, and necrophile: of taking on that *character-role*. To say that he is a rapist (etc.) is true *of* him within the fiction even though it is not something that is necessarily true *about* him (recall discussion in chapter 2); nor is it necessarily true that he desires to be such a person for real. Instead, S_1 enjoys$_{(sim)}$ doing these things, where what "these things" refers to are fictions he is depicting by singing along. His game world is not, in and of itself, immoral and so the fun/thrill he derives from this is not immoral. Nevertheless, in accordance with Oswell's point presented at the end of the last chapter—that "the ethical intensity of the virtual [or fictional] image lies precisely in its capacity to refer to a scene beyond itself" (2006, p. 258)—one may hold that what S_1 is doing is *trivializing* the immoral acts the fictions depict. In this case, for his own enjoyment. As such, one may feel justified in accusing him of *poor taste* (see chapter 9 for further discussion on this point).

S_2, like S_1, enjoys singing along to *Mind of a Lunatic*. Within S_2's game world, however, he fictively imagines that the immoral acts are real in order for the fictional content (his game world, derived from the song's work world) to act as a vehicle for his actual desire. His enjoyment is therefore akin to enjoyment$_{(sub)}$, making the game world he has constructed immoral and his enjoyment$_{(sub)}$ of the fiction likewise immoral. To be clear, the game world content is not immoral per se (in accordance with the conclusion drawn in chapter 3); rather, the construction of the game world—its essence, if you like—is immoral because it is a fictional space created for the enjoyment of that which is immoral. It serves no other purpose (for S_2), and enjoying that which is immoral is itself immoral, in accordance with P_0.

CONCLUSION

Throughout the chapter, I have sought to show how it can be both immoral and not immoral to *enjoy* depictions of fictional immorality. Whether it is one or the other turns not on the content one is enjoying per se (which accords with the conclusion arrived at in chapter 3), at least not in terms of the content of the particular work world one is engaged with but, instead, on the type of enjoyment one is trying to procure and, as a consequence, the game world one creates as a means of procuring this enjoyment. Proposition $p_2{**}_{\text{REFINED}}$ serves as a sufficient condition for the immorality of enjoying fictional depictions of immorality because it is sensitive to the difference in enjoyment presented within the chapter between enjoyment$_{\text{(sub)}}$ and enjoyment$_{\text{(sim)}}$, as well as Walton's work world/game world distinction. As such, $p_2{}^{**}{}_{\text{REFINED}}$ permits the denouncement of S's act of enjoying$_{\text{(sub)}}$ depictions of fictional immorality, *qua* his game world, while simultaneously allowing that the depiction of fictional immorality (*qua* work world) from which S's game world is derived is not immoral. Conversely, the rejection of $p_2{}^{*}$ allows S to enjoy$_{\text{(sim)}}$ fictional-x_{i} without succumbing to a charge of immorality.

In contrast to one's enjoyment of fictional immorality, it may be that one is unwilling to engage with certain fictional content. Understanding why this is could offer insight into the sorts of depictions that ought to be permitted and those that ought not. In short, then, could our willingness or unwillingness to engage with certain fictions be co-opted as a measure of morality? In the next chapter, I examine the phenomenon of *imaginative resistance* and consider its utility as an indicator of moral wisdom.

NOTES

1. See Kershnar (2005), Moore (1966), and Ross (1930) for a similar view.
2. Enjoying the enactment for its own sake may be symptomatic of S's *aesthetic attitude* (Gaut, 2007), whereby what S enjoys is the aesthetic quality of the depiction in and of itself.
3. In Young (2013), I also present a third motivation that is perhaps best suited to a video game context, but need not be: namely, M$_{\text{(strategic)}}$. Based on this motivation, one engages with the fictional immorality because it benefits one's overall strategy, which is to win the game. M$_{\text{(strategic)}}$ is compatible with Ostritsch's weaker sense of enjoyment, although it is not the same.
4. The gameplay consists largely of stalking and sexually assaulting women (i.e., a mother and her teenage daughters).
5. Whether one considers the game world to be aligned with the authorized work world perhaps rests on whether one holds that an authorized game world should *not* include a desire for what one is make-believing to actually be true. I favor this interpretation.

Chapter 8

Resisting Fictional Immorality

INTRODUCTION

As well as enjoying depictions of fictional immorality, engaging with fictional-x_i can elicit from its audience strong negative emotions (as part of what Smuts, 2016, calls the imaginative experience). We may be sickened by the brutality inflicted on a character (recall the example of people fainting during Dickens' rendition of the murder of Nancy by Bill Sikes), be both angered and distressed by scenes of cruelty (e.g., the Thénardiers' treatment of Colette in *Les Misérables*), or be horrified by the actions of a deranged killer (e.g., Patrick Bateman from *American Psycho*), and so on. Yet, paradoxically, these same negative emotions, which we tend to shy away from in other aspects of our lives, are sometimes, perhaps even often (for some people, at least), one of the main motivations, if not *the* main motivation, for engaging with fiction (Smuts, 2007, 2014).

In the 2014 UK TV drama *The Missing*, for example, we are invited to bear witness to the anguish of the parents of a missing child as they desperately await news of his whereabouts and possible abduction. In particular, the fiction catalogues the plight of the father (played by James Nesbit), who holds himself responsible for what happened, and his slow and painful descent into madness as he searches in vain for his son. Sam Wollaston of *The Guardian* has this to say about *The Missing*: "This eight-part thriller is not mawkish, or overly sentimental; just very human, and very very sad. James Nesbit is so very very good at pain; he doesn't just share it, he forces it on you" (p. 1).[1] Likewise, we expect to be scared when engaged with horror fiction; indeed, we use a phrase like "it wasn't scary" as a criticism, not by way of relief. As such, not only do we expect to be scared; more than this, we hold in high regard examples of the genre that succeed in doing this (Carroll, 1990). Importantly, though, while each example illustrates those occasions when negative emotions are elicited from, and perhaps even sought by, the audience, my interest in this chapter is not in the *paradox of painful art* (as it is sometimes called). I am not concerned with why rational individuals seek out fiction in the expectation—or hope—that they will experience negative emotions. Nor do I intend to revisit the way in which a morally pertinent message is being conveyed by the fiction or the manner in which we are being invited to reflect on this: say, through depictions of cruelty or brutality or horror. I have discussed this issue already in chapter 4. Instead, I am interested in the psychological phenomenon of *imaginative resistance* and how it can be used to inform our moral judgments about depictions of fictional immorality.

Following the distinction I introduced in the last chapter between Walton's (1990) work worlds and game worlds, imaginative resistance is a means by which we indicate that there are some game worlds that we (or at least some of us) are unwilling to create through our fictive imagination because the work world depicts certain acts/events/relationships that run contrary to our own moral values. To be clear, if one engages in imaginative resistance, in order to resist creating a particular game world, then it is not that one lacks the ability to make-believe *p* (where *p* is a particular immoral depiction found within the authorized work world); rather, it is that one is unwilling to do so. I am therefore distinguishing imaginative resistance from hermeneutic recalibration (Liao, 2013) mentioned in chapter 4. Thus, while I have the cognitive capacity to fictively imagine that a future Earth exists in which the ideology of white supremacy is accepted as the political norm, and therefore where the truth of this ideology is presented as a truth of *this world*, I am unwilling to do so. The same applies to a fiction told from the perspective of the elites of an imaginary society who advocate the (make-belief) virtue of pedophilia as a form of training in servitude for all classes who do not comprise the social elite. To be clear, what I am describing, here, is not resistance to the act of fictively imagining (*qua* one's game world) a world that *depicts* racial discrimination or sexual assault from the perspective of one who seeks to challenge it, or from a fiction that provides the means to do this (say, by highlighting the wrongness of the act); rather, I am describing examples of fiction from the perspective of those, *from within the fiction*, who endorse it (although, from their perspective, they would not consider their actions to be wrong). What is being resisted, then, is the make-believe truth of the fiction's expression of truth as understood from within the fiction. So, the truth of white supremacy within the example given above is true only in the way that it is true that Sherlock Holmes resides at 221b Baker Street. Yet imaginative resistance holds that we are far more willing to make-believe the truth of the latter than the former.

The aim of this chapter is to examine the viability of co-opting imaginative resistance (*qua* our psychological response to the invitation to engage in certain game world fictive imaginings) as a moral marker, and therefore as a normative measure of fictional immorality. What is there to be gained from co-opting imaginative resistance in this way? I am not suggesting that the mere occurrence of imaginative resistance is, in and of itself, sufficient to justify the moral condemnation of depictions of fictional immorality (irrespective of how much intuitive appeal this may have for some). Instead, I seek to understand whether imaginative resistance has the potential to inform judgments about the morality of fictional-x_i, an implication of which is that on those occasions when an individual or group of likeminded individuals do not engage in imaginative resistance—in the context of fictional immorality—then we have the resources at hand to justify the claim that such resistance is nevertheless warranted. The potential utility for co-opting explanations of imaginative resistance is therefore high. Importantly, though, all of this is possible only *if* explanations of imaginative resistance can be co-opted in this way.

In what is to follow, I intend to show that imaginative resistance is a psychological aberration that lacks the resources to act as a sufficient measure of what it should be permissible to depict by way of fictional immorality.[2] By "aberration," I mean a deviation from an otherwise much more open willingness to fictively imagine fictional events. That said, I accept that any psychological constraint seems to act as a reliable indicator of personal taste (a point I shall return to at the end of this chapter and discuss in detail in the next). Therefore, exploring the phenomenon of imaginative resistance is far from futile, even if it is unable to contribute to a moral position with regard to depictions of fictional immorality that has moral import beyond an expression of personal taste.

As a reliable indicator of personal tolerance, imaginative resistance is able to provide information on what different people find *intolerabl*e to fictively imagine (*qua* their personal game world), and therefore what they are likely to judge is in poor taste and insensitive. Before discussing this in detail, however, I will begin by outlining the context in which imaginative resistance is said to arise.

IMAGINATIVE RESISTANCE: AN UNWILLINGNESS TO EMBRACE "IMMORAL" DEPICTIONS

Consider the following sentences:

1. James T. Kirk succeeded Christopher Pike as the captain of the USS Enterprise.

2. Ebenezer Scrooge once had a business partner called Jacob Marley.
3. A man on a lonesome road walks in fear and dread because he knows a frightful fiend does close behind him tread.[3]

Each sentence describes a fictional event or relationship. Recall, however, from chapter 2, that, according to the nonpredication theory of literature, each statement is neither true nor false because the presupposition of each statement is not taken to be true. In other words, the events/relationships described in (1)–(3) can be neither true nor false because that which would otherwise be taken for granted in the case of each of the characters described—the truth that they exist or existed—is understood not to be true. If it is understood that Kirk and Pike do not exist—have never existed and will never exist[4]—then the statement described in (1) lacks truth-aptness. Yet, from chapter 2, we also learned that there is a sense, at least within the fiction, in which it is *true of* Kirk that he succeeded Pike, or *true of* Marley that he was Scrooge's partner, and so on, even if neither of these things is *true about* Kirk or Marley, respectively (see also Alward, 2010).

According to Stern (1965), the truth of the information contained within statements (1)–(3) does not equate to a truth that is used to inform (or misinform, in the case of a falsehood); rather, it is used to *perform* the task of creating a fictional event/character/relation. The fictional narrative should therefore be viewed as a prop, which guides the subject in the construction of the make-believe (Mothersill, 2006) (or, in Walton's parlance, their game world). As a consequence, while it is neither true nor false "in fact" that Kirk succeeded Pike, one is nevertheless free, although required (if the fiction is to work as fiction), to perform the act of *make-believing* (through the creation of one's game world) the truth of Kirk's succession to the captaincy, or that Marley was Scrooge's former business partner, or that the fiend is close at hand (Everett, 2007; Walton, 1990). Thus, in cases where a proposition lacks truth-aptness, through an act of fictive imagination, it can nevertheless be make-believe true that *p* (see also, Stern, 1967).

Within each of the fictions described in (1)–(3), not only is it possible to make-believe their respective truth but, as a matter of routine, it is something we are willing to do. Imagine, then, that the fiction in (1)—concerning the captaincy of the USS Enterprise—continued thus: Other candidates were not considered for the position of captain because, of those available, all were homosexual, and it was right that they were excluded because homosexuality is morally repugnant and no starship captain should be associated with that which is morally repugnant. Or, in (2), that Jacob Marley engaged in nonconsenting sexual intercourse with women in front of an approving crowd, and it was acceptable to do so because they were women. Or, finally, that the frightful fiend in (3) was a frightful fiend because he was a black man, whose fiendishness went hand-in-hand with his (and all other black people's) inferiority to whites.[5] In essence, what is being described directly or implied in each respective case is one of the following views:

(i) Homosexuality is morally wrong.
(ii) Engaging in nonconsenting sexual intercourse with women is morally right or at the very least not morally wrong.
(iii) Black people are fiends and inferior to whites.

Importantly, statements (i)–(iii) are derived from a fiction. Yet, according to proponents of imaginative resistance, we are far more willing to fictively imagine and therefore make-believe that Kirk succeeded Pike than we are that homosexuality is morally wrong (Gendler, 2000).[6] Or, put another way, within a particular fiction, we are far more willing to accept the *truth of* Kirk's succession than the *truth of* homosexuality's moral deviance. In the case of the former (nonmoral) utterance, the authority of the author to present us with a *fictional truth* proceeds unchallenged (Stueber, 2011), whereas with the latter depiction, and hence fictional truth, concerning the moral status of homosexuality, we are likely to experience what Gendler (2006) refers to as "pop-out": a sudden disengagement with the fiction. Fictional moral claims that run contrary to our own morality are therefore likely to be resisted and the authority of the author on such matters undermined (Nanay, 2010; Todd, 2009). This *authoritative breakdown* (Gendler, 2006) occurs far less frequently in cases of nonmoral fiction, even where fictional facts deviate radically from what is actually known (e.g., that Germany won World War II; see chapter 10 on historical fiction).[7] It is not that I might resist fictively imagining *qua* creating a game world in which some people within the fiction *believe* propositions (i)–(iii), and therefore that I might resist being drawn into the fiction through their eyes (as it were); rather, I might resist fictively imagining *qua* creating a game world in which the moral position *is* make-believe true (*true of* X), not just believed to be true by some of the characters within the fiction.

When watching *Star Wars*, for example, my engagement with the fiction does not take the form of believing that some of the characters believe that they are part of a Rebel Alliance, irrespective of whether this is the case or not, or believe they are onboard the Death Star. Rather, I make-believe that this is the case and accept the truth *of* these things. Yet, imaginative resistance indicates that I am less willing to adopt the same strategy with a fiction in which the truth of certain morally pertinent depictions goes against my own moral convictions. In such a situation, I am likely to resist creating a game world in which it is *make-believe true* that the fictional characters are *justified* in what they believe (e.g., that no starship captain should be homosexual), even in cases where what these characters believe reflects the moral norm within the fictional realm (Mothersill, 2006). In short, within a given fiction, imaginative resistance informs us that I am likely to be unwilling to accept the *truth of* X (within the fiction) if I am morally opposed to the same *truth about* X (beyond the fictional realm).

Proponents of imaginative resistance accept, somewhat intuitively it would seem, that we typically resist imagining the sort of deviant moral claims outlined in (1)–(3), and therefore resist creating a game world corresponding to the authorized work world of that particular fiction. Walton (2006), while acknowledging that assertions about imaginative resistance should be subject to empirical verification, nevertheless seems unsurprised by *the idea* that people show the kind of resistance discussed. Yet if I am willing to fictively imagine (as part of my game world) that there *is* a world where Kirk succeeded Pike as the captain of the USS Enterprise and not just fictively imagine that this is *believed* to be the case by some within the fiction, then should I not be equally willing to *fictively imagine* (rather than actually believe) that there is a world where homosexuality *is* morally repugnant, in some categorical sense, and not just *believed* to be so (again, by some within the fiction)? Equally willing, that is, to *make-believe* that categorical truths exist which denote the immorality of homosexuality or that engaging in nonconsenting sexual intercourse with women is morally good or that the color of one's skin determines one's moral character. After all, we are "not more convinced of the evil of slavery than we are of the non-existence of dragons and fairy godmothers" (Mothersill, 2006, p. 76); so why is it that we are more willing to imagine the existence of the latter than the moral goodness of the former? As Walton (1994) asks: If we embrace science-fiction, why not moral-fiction?

In summarizing this section, prima facie, imaginative resistance is a *psychological* measure of our unwillingness to engage with certain depictions of fiction: more often than not, depictions of fictional immorality. Therefore, given the context in which discussion on imaginative resistance tends to occur, it seems reasonable to explore the extent to which the phenomenon could be co-opted as a way of morally validating which, if any, depictions of fictional immorality one should not be willing to engage with. In order to understand why it might make a valid measure, we need to consider explanations for its occurrence. Understanding why resistance occurs might inform moral arguments against certain depictions in certain contexts or when certain conditions are, or seem to have been, satisfied.

Two possible explanations have already been touched on in previous chapters, although not in the context of imaginative resistance. These are ambiguous meaning (chapter 4) and risk of harm (chapter 5). It is worth broaching these topics again, however, in order to achieve a broader understanding, although I am mindful of unnecessary repetition and so will limit discussion to explaining imaginative resistance.

AMBIGUITY WHEN VIOLATING FIXED NORMS

When directed toward fiction, our moral beliefs appear to have a transcendent quality that is lacking (perhaps through willing abdication) in cases of nonmoral beliefs. Recall from chapter 4, the distinction Booth (1988) makes between nonce beliefs (which we are required to hold for the duration of the fiction but which are not intended to be exported from the fictional world) and fixed norms (which are meant to traverse these two worlds and provide the backdrop against which we judge the exploits of the protagonist), among others. The nonce beliefs are a contingent vehicle through which we convey "truth" about the fixed norm(s). Therefore, whether set within a zombie apocalypse, a galaxy far, far away, or closer to home, the wicked (for example) should not ultimately triumph.

It may be, of course, that the author, as part of the fictional narrative, taken as a whole, or through an isolated depiction, attempts to provoke the audience by subverting these fixed norms and so challenges certain traditional moral values. Where *p* is morally bad in our world, let us say that the fiction presents it as morally good or, at least, not morally bad (recall the example of Joseph Heller's *Catch 22*). Therefore, when invited to engage with a fiction in which the moral goodness of *p* is avowed, thereby subverting a moral norm that exists outside the fiction, we may resist engaging with it as part of our game world because we cannot be sure that the author, through the fiction and therefore through the work world, is not making simultaneous claims about the fictional and nonfictional worlds (Gendler, 2006). In other words, "when we encounter fictional truths that depict deviant morality, we cannot assume that their deviance is an indication that the author does not wish them to be exported" (Gendler, 2000, p. 78).[8]

It may be, of course, that the author, via the intentionally subverted depictions, is provoking us into taking a moral stance, thereby making his or her actions, along with fictional-x_i, defensible as a means of moral education (as discussed in chapter 4). But, equally, if, owing to ambiguity, we believe (or merely suspect) that the author's invitation to engage with fictional-x_i (*qua* the work world) is designed to manipulate us into rethinking our moral beliefs, then we may decide that we are no longer willing to be spectators of or even participants in the fiction, because we hold ourselves to be the victims of an attempted manipulation (Moran, 1994). I may, for example, come to believe (or merely suspect) that I am being manipulated into fictively imagining, and so make-believing, that *p* is good in order that I might eventually come to *believe* that *p* is good, rather than simply continue to make-believe this.[9]

Where it is not known that the fiction, taken as a whole, endorses an immoral worldview and therefore where it is not known that fictional-x_i is congruent with the endorsement of an immoral worldview (thereby failing to satisfy IWV; see chapter 4), but also where a reasonable interpretation of the fiction leaves unclear whether the fiction is endorsing or merely depicting an immoral worldview, then although one is not justified in claiming that fictional-x_i is *immoral* (recall from chapter 4 that $\text{IWV}_{\text{revised}}$ was rejected), one should nevertheless feel justified in refraining from engaging with the fiction further. This is because to engage, in the absence of mitigating circumstances, would be, minimally, morally insensitive (in accordance with $\text{WV}_{\text{insensitive}}$) and, as we shall discuss in the next chapter, perhaps even in poor taste.

It may be that imaginative resistance occurs because of a *belief* that fictional-x_i is immoral because it is endorsing the immoral. What I am claiming here, instead, following the discussion in chapter 4, is that any restraint characteristic of imaginative resistance is justified as a resistance to the immoral *if* the depiction satisfies IWV (sufficient condition). Where it does not, then such resistance may still be justified *if* it satisfies $\text{WV}_{\text{insensitive}}$, and is therefore expressing resistance to a fiction that is, minimally, morally insensitive.

THE RISK OF NEGATIVE EFFECT (HARM)

Where one is invited to engage with fictional-x_i, and therefore fictively imagine the truth of a morally deviant claim, irrespective of any ulterior motive on the part of the fiction's creator, or even in the full knowledge that their intention is benign, then imaginative resistance may still be a means of expressing our unease at the depiction, insofar as it may be symptomatic of the fact that we *fear* that it will lead to our becoming morally corrupt, such that we might come to believe and even delight in what we had previously only fictively imagined was true (McCormick, 2001). Now, engaging in a fiction (*qua* work world) and, in doing so, make-believing (*qua* one's game world) that *p* is good is unlikely to result in an *immediate* weakening of one's conviction that *p* is in fact morally bad; it may not weaken one's conviction at all, but I may fear that it will lead to a gradual change if I continue to engage with the fiction. Either way, the make-belief may unsettle me, perhaps causing moral "disorientation" (Walton, 1994) which I will likely find unpalatable, again causing me to resist engaging with fictional-x_i.

If someone feels uncomfortable engaging with a deviant moral-fiction, then, *ceteris paribus*, that is reason enough *for them* not to do it, and so reason enough to exhibit imaginative resistance. But, given the discussion on harm in chapter 6, in which there is inconclusive evidence to support the case for moral harm as a direct result of engaging with depictions of fictional immorality, it is not reason enough to claim that it should not be done, *qua* some form of overarching moral pronouncement. Consequently, it is not reason enough to prohibit, morally, those who are willing to engage with depictions of fictional immorality and therefore fictively imagine (*qua* one's game world) deviant moral truths.

VIOLATING THE SUPERVENIENCE RELATION

Gendler (2000) argues that moral claims are commonly believed to be categorical insofar as what is held to be morally good or bad in one world is equally so in all possible worlds. Consequently, if it is wrong to murder or rape or discriminate in one world, then this should be true of all worlds, including fictional ones. Supporting Gendler's explanation is the widely held view among moral philosophers that moral properties supervene on nonmoral properties (Blackburn, 1985; Depaul, 1987; Meyers, 2012). In other words, a moral property M (the "wrongness" of *x*, for example) is said to be *supervenient* on a physical (nonmoral) property P such that there can be no change in M without a change in P. Where A and B share the same nonmoral properties, A and B share the same moral properties. This supervenient relation is famously illustrated by Hare (1952):

> Suppose that we say "St. Francis was a good man". It is logically impossible to say this and to maintain at the same time that there might have been another man placed exactly in the same circumstances as St. Francis, and who behaved in exactly the same way, but who differed from St. Francis in this respect only, that he was not a good man. (p. 145)

The relationship between the moral property of "goodness" and the nonmoral or physical properties that constitute St Francis—his physical description—is such that there can be no moral change in the absence of a physical/behavioral change. It therefore seems to be the case that if it is said of A that A is morally good, then, a priori, where A is identical to B, this entails that it is likewise said of B that B is morally good. A failure to conclude that B is morally good indicates a failure to understand moral concepts. A consequence of this a priori truth is that the moral property of "goodness"—which, in this case, renders St. Francis a good man—is dependent on some physical (nonmoral) property of the right kind. But "moral goodness" is not reducible to this physical property; it is not something that we can describe in purely physical terms. As such, in addition to there being no change in M without a change in P, moral supervenience holds that M is irreducible to P (Ridge, 2007). Neither is the relationship analytic, whereby a physical description of A (or B, if identical to A) would reveal to us those components that make up moral goodness. In other words, the proposition "A is morally good" is not analytic in the way "a triangle has three sides" is. Furthermore, we do not directly detect, a posteriori, the actual instantiation of a moral property (say, wrongness); rather, we "conclude that particular acts are wrong in virtue of some empirically detectable feature; for example, because it *causes pain*, *involves deceit*, or *violates an agreement* (etc.)" (Coons, 2011, p. 85; emphasis in original) and, in doing so "postulate the unobserved to explain the observed" (*ibid.*).

This, of course, does not mean that *any* physical difference necessitates a moral difference. If a physical duplicate of St. Francis were suddenly to find that his hair had gone gray overnight or that he had suddenly become short-sighted, then it is unlikely (or certainly would not necessitate) that this would have any impact (detrimental or otherwise) on his "goodness." Any physical difference between the duplicate saints, if it is to be cited as the reason for a moral difference between the two men, must constitute *reasonable grounds* for this difference (Sidgwick, [1874] 1981); it must be a physical difference of the right kind. As an aside, the precise nature of the supervenient relation is also dependent on whether it is held to be local/global, weak/strong (Kim, 1993). This difference is not of major concern here, however, although I will touch on it briefly below.

Walton (1994) argues that imaginative resistance in the face of deviant fictional morality occurs because our required make-belief (*qua* game world) about such morality contravenes the supervenience relation (see also Weatherson, 2004). Where the property of moral goodness supervenes on, say, the nonmoral act of generosity, if A and B are identically generous, then A and B should be judged equally good. If we are told in the fictional world that A is generous, which matches our understanding of the term "generous" when used outside of the fiction, then it should follow that A is good, given the supervenience relation just described. Should A's generosity be depicted as bad, however, then given what we understand about generosity and its relation to goodness, the fictional claim that A is bad should appear incomprehensible to us. It should be difficult for us to understand how two identical actions—one from the fictional world and the other nonfictional—can be described as good in one world but bad in the other (Gendler, 2006). Likewise, if, in another fiction, sexual assault by the social elite on the nonelite is depicted as morally good, because it creates subservience in the nonelite classes, then, again, this may appear incomprehensible to us. In either case, we would likely be reluctant to engage with the fiction's moral depiction.

In short, for fictional ϕ to be morally good, where nonfictional ϕ is morally bad, the concept of supervenience requires that the fictional and nonfictional descriptions of ϕ differ in some way that constitutes reasonable grounds for this difference. Where they do not differ, logic dictates that the moral status of ϕ must be the same. Of concern, however, is not whether ϕ is morally bad in some moral realist sense that *necessarily* transcends worlds owing to identical physical descriptions of nonmoral properties of the right kind—I am, to be clear, an antirealist when it comes to morality (a point I will take up in greater detail in chapter 13)—but whether I, in accordance with a societal moral norm, *believe* that ϕ is morally bad.

From an antirealist perspective, when spoken intelligently and sincerely, the proposition "A is morally good" is equivalent to asserting "I *believe* that A is morally good." In light of this, and given that the relationship between A and moral goodness is not analytic, nor is moral goodness directly detectable a posteriori, it is left unclear how our moral beliefs are justified. (Again, I will have more to say on this in chapter 13.) For now, however, a way to address this problem is to hold that the proposition "A is morally good" is not expressing a belief about some mind-independent moral fact attributed to A, which can be either true or false, but, rather, one's attitude to A that, broadly construed, amounts either to one's approval or disapproval of A.[10] Importantly, though, the supervenient constraint should still hold in the case of one's moral *attitude*. The proposition "ϕ is morally wrong," for example, should be thought of as a direct *expression* of one's attitude to ϕ, which functions in an evaluative way. Consequently, where A is identical to B, if one morally approves of A (*qua* has a positive attitude toward A), then one should morally approve of B, also.

Given the supervenient relation that exists between ϕ and one's attitude to ϕ, in accordance with this, one can accept that, where A and B share the right kind of nonmoral properties, what one *believes* about the moral status of one must be consistent with one's belief about the moral status of the other. If such a position were tenable, then, it could explain imaginative resistance to fictional depictions of moral violations. But even if it could explain imaginative resistance, accounting for a psychological phenomenon does not justify using this as a measure of morality in the case of depictions of fictional immorality. The extrapolation from psychological phenomenon to normative position is a problem in and of itself. In addition, however, the supervenience relation is not violated, as I shall now explain.

It is not true that x_i and fictional-x_i—say, sexual predation and a depiction of sexual predation—share all the same nonmoral properties *of the right kind*, even if what is descriptively true *of* x_i is descriptively true of fictional-x_i. To borrow an example from chapter 2, the description "John Smith is a sexual predator" is equally true *of* an actual sexual predator called John Smith as it is *of* a fictional character with the same name depicted doing similar things. Importantly, though, there is something *about* x_i and fictional-x_i—a property of the right kind—that is different, and it is my contention that this difference prevents a violation of the supervenience relation in cases where (for example) x_i is morally bad but is nevertheless depicted as being morally good in *a* particular fiction. A truth *about* x_i and fictional-x_i that is not shared is that the latter is not truth-apt or, put differently, does not have the property of truth-aptness, whereas the former does. The fact that fictional-x_i lacks truth-aptness reflects my altered epistemic relation to it (compared to nonfictional x_i): in the former case, I only fictively imagine *qua* make-believe its existence and, if applicable, its moral inversion (i.e., being depicted a morally good rather than bad). The differing property of truth-aptness evident in the case of fictional and nonfictional x_i, and the altered epistemic relation that subsequently follows, means that I can believe that x_i is morally bad while make-believing its moral goodness (in virtue of the depiction I engage with) without violating the concept of supervenience (we have, it would seem, reasonable grounds for a difference, as Sidgwick might say).[11] Moreover, in the case of fictive imagining, it is my contention that the supervenient relation is sufficiently local and weak (Kim, 1993) to require only "same world" consistency in make-belief. If, within the same fiction (and therefore fictive imagining), two identical descriptions of ϕ are presented, then the make-believe truth of the same moral status must supervene on ϕ. Where two separate fictions describe ϕ in the same way, the make-believe truth of the moral status of ϕ need not be consistent across these fictional worlds. One is free to make-believe that ϕ is good in one fiction and bad in the other without violating the (weak and local) supervenience relation.

Whether allegedly violating the supervenience relation is a satisfactory explanation of imaginative resistance as a *psychological phenomenon* is not of concern here. Even if (for the sake of argument) it is—insofar as resistance is caused by the fact that one (mistakenly) *believes* that x_i and fictional-x_i share the same nonmoral properties of the right kind to warrant the same moral properties—closer inspection of fictional and nonfictional x_i reveals a pertinent difference between the two that in fact permits a moral difference without violating the supervenience relation, even when the same descriptive properties are true of each. This nonmoral difference concerns not what is true *of* x_i and fictional-x_i but what is true *about* them (the former is truth-apt; the latter is not). Irrespective of its worth as an explanation of the underlying psychology, then, in relation to identifying a means of measuring the morality of depictions of fictional immorality, violating the supervenience relation fails. Perhaps this is for the best, as the idea that I should be morally condemned for engaging with a fiction that depicts (inter alia) the moral worth of killing an innocent person or forced sexual intercourse or torture, simply because doing so violates the concept of supervenience (which, as it happens, it does not), seems to understate somewhat the moral controversy involved in these examples of fictional immorality.

CONCLUSION

As we reach the end of the chapter, what I hope to have shown is support for the assertion made at the start: that imaginative resistance does not amount to a sufficient means of measuring the morality of depictions of fictional immorality. While it may be the case that imaginative resistance is symptomatic of our (or at least some people's) intuitions/fears about the fiction's (and therefore the creator's) ulterior motives, or the potential negative effects of engaging with *these* depictions of immorality (whatever they happen to be), in the end, each is reliant on empirical confirmation (i.e., that there is an ulterior motive or that the potential risks are real). These possibilities have been discussed at length in previous chapters and were revisited briefly, here. In each case, their utility as a measure of the immorality of fictional-x_i is dependent on evidence and remains an open and testable empirical question.

I propose that what imaginative resistance does indicate, for those who exhibit the phenomenon, is a measure of personal tolerance of depictions of fictional immorality based on one's own degree of moral sensitivity and perhaps beliefs about morality. While this is not reason enough to bestow on the phenomenon the status of objective measure of morality, for the reasons discussed, it is nevertheless reasonable, I feel, to acknowledge its potential as a reliable indicator of *personal* morality or, minimally, those depictions of fictional immorality one finds unpalatable, even distressing. Moreover, as (potentially) a reliable indicator of moral sensitivity, I would anticipate a strong correlation between imaginative resistance and depictions of fictional immorality considered to be in poor taste. The matter of taste is the focus of the next chapter.

NOTES

1. https://www.theguardian.com/tv-and-radio/2014/oct/29/the-missing-james-nesbitt-review (accessed May 20, 2019).

2. This chapter was originally published as Young, G. (2015). "Are there some things it is morally wrong to make-believe? An examination of imaginative resistance as a measure of the morality of pretence." *Cogent Arts & Humanities*, 2(1), 1076956. https://doi.org/10.1080/23311983.2015.1076956"

3. Adapted from *The Rime of the Ancient Mariner*, Part VI.

4. Even if persons named Kirk and Pike existed in the past or will exist in the future, it is understood that these persons are not being referred to here (see chapter 2).

5. Examples borrowed from Young (2014a).

6. Weinberg and Meskin (2006) distinguish between imaginative resistance as *imaginative refusal* and *imaginative blockage*: the former equates to an unwillingness to imagine, the latter an inability to imagine. It is imaginative resistance as refusal

rather than blockage that is of interest here. Similarly, Gendler (2006) distinguishes between imaginative resistance as *imaginative barriers* (similar to imaginative blockage) and *imaginative impropriety*. It may be that my unwillingness or refusal to imagine "that *p*" is because I consider doing so to be improper.

7. Our willingness to imagine nonmoral deviation is not totally unconstrained. Yablo (2002) and Weatherson (2004) each discuss examples of nonmoral deviation that encounters resistance: finding a five-fingered maple that is oval shaped, for example, or a television and armchair that are indistinguishable from a knife and folk.

8. Such a position contrasts with Eagle (2007), who maintains that what we learn about fiction is that the author is not trying to convey the literal truth. It may be, then, that imaginative resistance denotes that this lesson has not been learned.

9. Stueber (2011) argues that what we believe and what we are required to make-believe regarding a given moral position may appear incompatible simply because we are not provided with sufficient information about the reasons for the fictional moral truth. We may resist fictively imagining *p*, even if we do not perceive authorial manipulation, simply because the fiction's incomplete narrative results in a sense of incoherence. The failure to articulate the reasons for the moral position adopted within the fiction makes it difficult for the reader to empathize sufficiently (if at all) with the protagonist when acting in a way that contravenes our moral values. The lack of detail makes it difficult for us to understand why we should put aside our own moral values in this instance or at least quarantine our requested make-belief from them (as Stueber puts it) when imagining such deviant morality. It is not that we do not understand that someone (the protagonist in the fiction) may believe *p* (where *p* contravenes an accepted moral standard) and even have a reason for this belief; rather, it is that we do not understand—perhaps owing to the limited information available—how the protagonist's reason(s) could be considered within the fiction to be cogent and therefore justified. We resist the make-believe, Stock (2005) argues, because we fail to understand the protagonist's claim "from the inside."

10. For the sake of simplicity, I have discounted a morally neutral attitude.

11. The lack of true-aptness provides reasonable grounds for a difference in the epistemic relation with the moral property (e.g., *believing x* is morally good), irrespective of moral content. This means that I could believe that *p* is morally good while also make-believing that fictional *p* is morally good, or believe that *p* is good while make-believing that it is bad. The fact that there is a difference in the nonmoral description/properties of *p* and fictional *p* (truth-aptness) means that there is a difference between the nonmoral description and my *epistemic relation* to the moral properties that supervene on *p*: belief in the former case and make-belief in the latter. What I choose to make-believe can be the same or different in terms of moral content to what I choose to believe. The important difference is not the moral property but my epistemic relation to it.

Chapter 9

Poor Taste and Fictional Immorality

INTRODUCTION

In chapter 4, and in accordance with IWV, I argued that fictional-x_i is immoral *if* it depicts immorality congruent with an immoral worldview endorsed by the fiction. I also considered whether, in the absence of an unequivocal answer to the question of endorsement, fictional-x_i could nevertheless be charged with immorality based on a reasonable interpretation of the fiction's meaning. I concluded that the "reasonable interpretation" approach adopted by $\text{IWV}_{\text{revised}}$ was not sufficient to warrant a charge of *immorality* but that a case could be made for moral insensitivity, at least in the absence of mitigating circumstances. Recall that $\text{WV}_{\text{insensitive}}$ denotes a sufficient condition for moral insensitivity:

> $\text{WV}_{\text{insensitive}}$: Where a reasonable interpretation of a fiction, taken as a whole, cannot establish whether the fiction is merely depicting rather than endorsing an immoral worldview, then any depiction of immorality from within the narrative congruent with that worldview is, minimally, morally insensitive.

What I aim to show is that being morally insensitive satisfies the requirements of a suberogatory act (see section "Poor Taste as Suberogatory Action"). Before proceeding, however, it is important to note that a charge of moral insensitivity need not be directed at a depiction of fictional immorality solely because the meaning of the fiction—including an alleged endorsement of an immoral worldview—is ambiguous. If an audience does not believe that fictional-x_i is congruent with an immoral worldview, it may still hold that the depiction is morally insensitive, only for a different reason. Before discussing suberogatory wrongdoing, specifically, however, of immediate interest is the claim that depictions of fictional immorality are morally insensitive if the audience believes them to be in poor taste.[1] Thus, in this chapter, I consider:

(i) What makes a particular instance of fictional immorality something that is charged with being in poor taste

This is followed by an examination of:

(ii) The moral and hence prohibitive status of poor taste, including its status as a suberogatory act (alongside moral insensitivity more generally)
(iii) The role of aesthetics in differentiating between good and bad poor taste

Before continuing, it is worth noting that even though I accept that reference to *taste* is metaphorical, it is nevertheless my intention to treat the metaphor as a measure of wrongdoing that has a *moral* rather than aesthetic flavor (i.e., poor *moral* taste), albeit a flavor that has yet to be more clearly defined. I will therefore argue that judgments about poor taste reflect one's negative attitude toward the *treatment* of a morally pertinent matter (Young, 2019). As such, a depiction of fictional immorality is in poor taste *if* it is believed to *trivialize* (for example) what the depiction is a depiction of: inter alia racism, sexism, sexual assault, torture. As a point of clarification, trivializing a *morally pertinent matter* means that something can be in poor taste even if what is being trivialized is a morally good thing, such as famine relief, aid workers, or organizations like *The Samaritans*.[2] Given that my interest is in depictions of fictional immorality, however, I will focus on poor taste in relation to these rather than morally pertinent matters more broadly. Finally, this chapter should be viewed as the precursor to a wider discussion on differences between poor taste and immorality, which is taken up in chapter 13.

POOR TASTE AND OFFENSE

It is my contention that the sine qua non of poor taste is offensiveness. If S judges x to be in poor taste, then she will be offended by x; and if S finds x offensive, then she will experience, and likely express, some degree of discomfort and disgust toward x, including a desire to avoid x (Shimp & Stuart, 2004). To say that offense is the sine qua non of poor taste does not, however, restrict offense exclusively to the domain of poor taste. After all, one might find someone's insulting outburst both vulgar and offensive without judging it to be in poor taste (Archard, 2014). What is important is that where one does consider fictional-x_i to be in poor taste, one must do so because one believes that fictional-x_i realizes a property of which one disapproves *qua* finds offensive. Let us call this property O. More formally: if S holds that fictional-x_i is in poor taste, then S disapproves of O (*qua* offense) and believes that fictional-x_i realizes O.

In accordance with criteria set out in Archard (2014), for fictional-x_i to (be believed to) realize O:

(i) *Fictional-x_i* must be publicly observable
(ii) S must derive meaning from fictional-x_i such that fictional-x_i is perceived to be/interpreted as (inter alia) ridiculing or trivializing or showing irreverence to an individual or a group (e.g., victims of sexual assault). S must therefore believe that what is being communicated by fictional-x_i is directed at either them or some other individual, specifically, or a group they identify with, or be targeting some other group they recognize but do not belong to (i.e., a heterosexual finding a joke about "gay bashing" offensive).

In addition:

(iii) The context in which fictional-x_i is situated will likely contribute to whether it is believed to realize O and therefore be deemed offensive. To illustrate: While there is nothing implicitly offensive about a *depiction* of a generic murder victim, the winning entry at the 2016 Cornwall Beach Games sandcastle competition (in the UK) was a sculpture of a naked woman lying face down with a spade in her back, surrounded by police tape. The depiction of the fictitious murder victim was created by a team of local police officers. Despite winning the competition, local Falmouth councilor Hanna Toms reportedly said that the sculpture was in poor taste, although some members of the public were equally of the view that it was a lighthearted joke and therefore a bit of fun (a point I shall return to). The police later apologized for any offense caused (Evans, 2016).

It is worth stressing that criteria (i)–(iii) set out what is required for fictional-x_i to be considered offensive. This may not always be offensiveness characteristic of poor taste (the specific nature of which will become apparent as we progress). Nevertheless, if offensiveness is the sine qua non of poor taste, then offensiveness concomitant with poor taste must satisfy these criteria. Where they are satisfied, and one finds fictional-x_i to be in poor taste in virtue of the belief that fictional-x_i realizes O, then one is not describing an inherent property of fictional-x_i (i.e., its offensiveness); rather, one is expressing a negative attitude toward fictional-x_i.

In a study looking into factors that contributed to, or mitigated, the perception of offense at fictional depictions of violence (but not poor taste per se), Coyne et al. (2016) found that offense was taken when the person offended did not believe the fictional content was aligned with their own *values* or *codes of appropriateness* (e.g., perceived inhumanity or lack of empathy, or where the depiction was found to be overly graphic/gory or gratuitous). They also reported that women tended to be more offended by violent images than men, followed by those scoring higher on religiosity. Factors reducing the amount of offense taken, in contrast, were the person's perceived desensitization to violence, the unrealistic nature of the violence depicted, and (albeit to a lesser degree) the belief that violence is part of life and what is depicted reflects this.

Whatever their reason for being offended, those in Coyne et al.'s study, like those who were offended by the sand sculpture of the murder victim (e.g., Councilor Hanna Toms) disapproved of some property, O, and believed that the depiction—this particular fictional-x_i—realized a token of O. Those who felt the winning sandcastle entry was a bit of fun, however, like those not offended by *x* in the Coyne et al. study, did not believe that fictional-x_i realized a property of which they disapproved *qua* offensiveness.[3]

What tokens of O have in common—which identifies them as tokens of O and therefore as offenses characteristic of poor taste—is how they are construed as *treating* actions/events/relationships already identified within a society as immoral. A sufficient condition for fictional-x_i to be in poor taste is therefore:

> PT: Fictional-x_i is said to be in poor taste *if* one disapproves of (*qua* is offended by) O (*qua* a perceived treatment of immorality) and believes that fictional-x_i realizes O (treats the aforementioned immorality in a way that one finds offensive).

Recall from chapter 4 the advertisement for liquid soap depicting a fictitious murderer's blood-stained arm beside his victim. In accordance with PT, for S to be offended by fictional-x_i (the depiction of murder *in the advertisement*), and therefore consider fictional-x_i to be in poor taste, S must believe that it is treating something that S already finds immoral—namely, murder—in a way that she finds offensive. Let us allow that S finds the advertisement to be in poor taste because she believes that it is trivializing (the immoral act of) murder. Thus:

- S judges the advertisement for liquid soap to be in poor taste *if* (and therefore *because)* S disapproves of (*qua* is offended by) O (in this case, trivializing murder) and believes that the advertisement for liquid soap (fictional-x_i) realizes O (is treating murder in a way S finds offensive: i.e., trivializing it by treating it as a humorous means of selling a product).[4]

Moreover, where S believes that murder (x_m) is immoral, then whatever reason S has for believing this—that is, whatever moral system S adopts: be it the belief that x_m violates God's law or constitutes a failure in one's secular duty, or because it is a vice or increases negative utility, and so on (see, also, chapter 13)—the question of whether an act of *murder* has actually occurred is less ambiguous than whether one is treating murder in a trivial way. As such, it is much more likely that this particular x (murder) is believed to be immoral, not only because of a shared belief that murder is wrong but because this particular act (this x_m) can be and often is interpreted more readily as an act of murder. The same cannot always be said for cases of alleged poor taste, however, not necessarily because people vary with regard to what constitutes poor taste (although they may well do), thereby making a social norm regarding poor taste more difficult to construct, but because, even where such a norm exists, the particular example (say, fictional-x_i *qua* the murder in the advertisement for liquid soap) may be more ambiguous and therefore open to interpretation with regard to whether it is believed by those who make up a given society to realize O, and therefore whether it conforms to the common (attitudinal) point of view or social norm regarding taste (McAteer, 2016). (Recall, the sandcastle sculpture depicting the fictitious murder victim, and the fact that some onlookers viewed it as a harmless joke, while others decried it as poor taste.) Consequently, while it may be much easier to achieve consensus in the case of "*this* x_m is immoral," especially where x_m is unambiguously identified as such (i.e., in the case of premeditated murder), consensus is less easily achieved in (many) cases of poor taste.

Let us apply PT to a further example. In the original version of *Playing History 2: Slave Trade*, it was possible to engage a slave Tetris mode, which allowed "players to stack dead-eyed African bodies that have been squished into uncomfortable Tetris shapes into a slave ship" (Machkovech, 2015, p. 1). After protests, the Tetris mode was removed from the game. In accordance with PT:

- P (protesters) judged the Tetris *Slave Trade* game to be in poor taste *if* (and therefore *because*) P disapproved of (*qua* was offended by) O (the perceived trivializing of slavery) and believed that the Tetris *Slave Trade* game (fictional-x_i) realized O (was treating slavery in a way P found offensive: that is, was trivializing it by incorporating it within a game that reinforced the dehumanizing approach of the historical slave traders).

Of course, in keeping with the subjective nature of judgments about poor taste, and the problem of agreeing on a correct interpretation of an alleged case, the designers of the Tetris *Slave Trade* game defended their original decision to include the now-removed mode of play by stating that they intended for the depiction of Africans packed into a slave ship, in a manner akin to solving a Tetris puzzle, to be a means of conveying the real horror of the slave trade. They claimed, therefore, not to be trivializing the immorality and cruelty of slavery but to have found a novel way of illustrating it. No doubt most, if not all, of us will have a view on whether this example counts as poor taste or whether the intent of the designers, once understood, mitigates the charge (in accordance with discussion in chapter 4; see also the section "Bad 'Bad Taste' and Good 'Good Taste'").

THE NATURE AND STRENGTH OF "OUGHT" IN ACCUSATIONS OF POOR TASTE

To declare that ϕ is in poor taste is to intimate that ϕ ought not to be done. But this would also be the case if one were to announce that ϕ is immoral. If declaring that something is immoral or in poor taste intimates the same nominal outcome—namely, that it *ought* not to be done—then does this mean that accusations of poor taste and immorality convey the same strength of moral prohibition? I do not believe so. To understand why not, consider the following:

- If you wish to arrive at a particular destination more quickly, then you ought to take the high road rather than the low road.
- If you enjoy horror literature, then you ought to read Clive Barker.[5]

To say that one ought to do (or not do) x is to engage in a *directive speech act* (Searle, 1983). In both examples, "ought" is discretionary and equates to a *recommendation.* One is not obliged to heed the recommendation. Given this, recall the claim that:

(a) Pronouncing that "ϕ is immoral" intimates that ϕ ought not to be done

In (a), as with the two previous examples of a directive speech act, one is being directed. Only, now, one is morally called upon not to do ϕ (Ridge, 2014). Such a declaration goes beyond mere recommendation. Instead, one is required or obliged not to do ϕ. The "ought" in (a) is categorical; it conveys what morality *requires* (Silk, 2014; Sinnott-Armstrong, 1987). But what about:

(b) Pronouncing that "ϕ is in poor taste" intimates that ϕ ought not to be done

In (b), as with (a), there is a sense in which one is being morally called upon not to do ϕ; only, now, the "ought" seems to lie somewhere between a recommendation and a requirement. Taking the difference between (a) and (b) into account, we get:

(a*) Pronouncing that "ϕ is immoral" intimates that one is required or obliged not to do ϕ.
(b*) Pronouncing that "ϕ is in poor taste" intimates that one "ought" (in some yet-to-be-determined sense, lying somewhere between recommendation and requirement) not to do ϕ.

Where ϕ is held to be immoral, then not doing ϕ is a requirement, such that one is obliged not to do it.[6] But while one is obliged not to murder, one is not *obliged* not to trivialize murder, even if trivializing murder may evoke a sense of wrongdoing. Yet, to say that one ought not to do *x* because it is in poor taste seems to do more (as a directive speech act) than merely *recommend* to the perpetrator of the (alleged) poor taste that they desist. My decision not to take the high road rather than the low road, or not to read Clive Barker, despite my proclivity for horror fiction, is discretionary, and recognized as such. The same discretionary element seems less warranted in the case of events/actions judged to be in poor taste, however, or at least one would appear more socially constrained when it comes to how far one can exercise this discretion.

POOR TASTE AS A SUBEROGATORY ACTION

If I had to posit a "best fit" for the "ought" associated with poor (moral) taste, it is my contention that it is comparable to that found in a category of action Driver (1992) calls *suberogatory*. To illustrate, imagine three people visiting a cinema in which there are only three seats remaining: two seats together and a single. The three people comprise a couple and a single person. The single person, S, enters first, knowing the couple are behind. S takes one of the seats that are together, meaning that the couple will have to sit apart (which they would prefer not to do). S is aware of this but does not wish to move. In this example, S is within his rights to sit in any of the available seats and is therefore not obliged to move. He has done nothing that is morally prohibited; yet it seems that S's action is in some sense wrong. Similarly, consider the example of an able-bodied youth seated on a crowded bus who, on noticing a much older person forced to stand, fails to offer him her seat, even though the two of them could easily have swapped places. While neither is an example of poor taste, my point is that one might reasonably conjecture that the sense of poor taste I have in mind fits within the category of “blameworthy action” Driver has identified as suberogatory.

Neither of the previous examples is an example of poor taste because neither concerns the *treatment* of a morally pertinent matter. Instead, the way the lone cinema-goer and the youth acted toward the other people—their respective treatment of them—*is* the morally pertinent matter. What each agent did was in some sense wrong because they acted in an impolite or inconsiderate manner toward the person(s) concerned, even though neither agent was *morally obliged* to act in any way other than the way they did. Breaches of etiquette, impolite and/or inconsiderate behavior, all fall within the broad church that is suberogatory action.

To illustrate further, while it may be the case that overshadowing the bride on her wedding day—say, by wearing a more lavish and eye-catching outfit—is seen as a breach of etiquette, it is again not an example of poor taste because, as with the previous examples, the breach of etiquette concerns how one has acted toward (or, if you like, treated) another person directly, where the treatment itself is and has therefore created the morally pertinent matter, even where one is not (in this case) morally obliged *not* to undermine the prestige of the bride. Contrast this with the actions of a different wedding guest who, in an attempt to be humorous, remarks that there is at least one positive to come from the recent death of the bride’s mother at the hands of a drunk driver: the groom need no longer worry about incurring the wrath of the mother-in-law. Here, an accusation of poor taste could be leveled at S because of how he is perceived to be treating an already-existing morally pertinent matter: the unlawful and immoral death of the bride’s mother. To be clear, one could also argue that S was being inconsiderate toward the bride (and likely others) by telling the joke, and even that he showed a lack of respect to the memory of the bride’s mother. My point is that while there may be more than one layer or level to this wrongful act, the poor taste element stems from the treatment of a morally pertinent matter: namely, the mother of the bride’s untimely and unlawful death. This remains the case regardless of any other elements of wrongfulness one is able to associate with the telling of the joke in the context in which it occurred.[7]

This last example illustrates well the moral insensitivity evident within expressions of poor taste. Recall that moral insensitivity refers to one’s failure to show an awareness of—*qua* recognize and *appreciate*—the feelings and needs of others during (for example) one’s interpersonal interactions with them. Of course, one can and necessarily does exhibit moral insensitivity when engaged in any number of immoral activities; on such occasions, however, the moral insensitivity is somewhat consumed by the immorality of the act itself. Our interest, here, is where moral insensitivity exists in the absence of immorality. On such an occasion, one can be said to have engaged in an action that qualifies as a suberogatory-level wrongdoing.

BAD "BAD TASTE" AND GOOD "BAD TASTE"

John Waters, director of the 1972 film *Pink Flamingos* (which was advertised with the tag line: "An exercise in poor taste"), once said: "If someone vomits [on account of] watching one of my films, it's like getting a standing ovation" (Waters, 2005). He then goes on to draw a distinction between good bad taste and bad bad taste, perhaps as a means of defending his own work. (Note: When discussing Waters, I will use the term "bad taste," as he does, rather than "poor taste," for consistency of language.) I thus infer from Waters' second comment that he considers *his* expressions of bad taste to be good bad taste. In conjunction with his first comment, we can deduce that he believes that good bad taste is *capable* of eliciting a strong visceral response. If good bad taste is capable of doing this, and would be taken as a compliment by Waters if it did, even to the point of causing someone to vomit, then it seems reasonable to infer that bad bad taste is likewise capable of eliciting an extreme disgust response.

What determines whether bad taste is good or bad, and how might this inform our discussion on the matter of taste? Waters (2005) provides us with a hypothetical example of *bad* bad taste: a ninety-minute film of people getting their limbs hacked off. Leaving aside, for the moment, the question of what makes this an example of bad taste per se, for Waters, it amounts to *bad* bad taste because the film would be neither stylish nor original. The film's lack of style and originality is what makes it an example of *bad* bad taste rather than bad taste per se. It is unclear, however, whether Waters believes such a film's bad taste would be inherently bad or whether it could be transformed into a *good* bad taste film with an infusion of style and originality. I am going to assume the latter but, either way, my point is this: Waters seems to be basing his judgment about the goodness or badness of bad taste on its alleged aesthetic rather than moral quality. Thus, something that is in bad taste but is stylish and original, we can infer, would amount to good bad taste.

Waters (2005) continue thus: "To understand bad taste, one must have very good taste." One (but not the only) way to interpret Waters, here, is to hold that he is equivocating: by using the term "taste" in two different ways. The bad taste he claims we must understand (first use of "taste") is, I contend, reference to bad taste in relation to the treatment of a morally pertinent matter. To understand this type of bad taste, in a manner that allows one to depict it as *good* bad taste, one must possess good aesthetic taste (second use of "taste"): for it is through the application of one's (very) good aesthetic taste that the depiction of bad moral taste is transformed from bad bad taste to good bad taste. (Note: It is beyond the scope of this chapter to examine what enables one to possess good/bad/indifferent aesthetic taste as some kind of objectified [*qua* intersubjective] standard. For discussion on this in relation to Hume's notion of an expert judge, see Shelley, 1998, 2013; Williams, 2007.)

In order for one's aesthetic taste to mitigate bad taste (*qua* the treatment of a morally pertinent matter), one must be of the opinion that aesthetic appreciation can trump morally contentious matters of taste. Creating a fiction that is an example of good bad taste does not necessarily prevent a strong aversive reactions to what is depicted, however. As Waters concedes, producing good bad taste can be "creatively nauseating." To illustrate: Consider the portrait of convicted child murderer Myra Hindley created by artist Marcus Harvey from the handprints of children. The portrait provoked protests when it was first exhibited as part of the *Sensations* exhibition at London's Royal Academy in 1997 (see Young, 2000). According to my account of poor (moral) taste, those who protested likely believed that *x*—*this* portrait of Myra Hindley—was in poor taste (realized a property they disapproved of) in virtue of its perceived negative (e.g., irreverent, trivializing, exploitative) treatment of a morally pertinent matter—that is, child murder—and lacked sufficient aesthetic quality to mitigate this fact. For some, however, the aesthetic quality of the work may have overridden and therefore mitigated any morally contentious aspect of the work, creating good bad taste (to borrow Waters' term). For others, they may not believe that the portrait realizes any property of which they disapprove (i.e., they do not believe it trivializes or exploits child murder). As such, they may not consider it to be in poor (moral) taste at all.

Returning to the ninety-minute limb-hacking example Waters uses to illustrate bad bad taste, hacking off someone's limb(s) is clearly a morally pertinent matter, even in the case of a film set in an army field hospital during some fictitious war where the surgeon is removing limbs to preserve life (causing a *pro tanto* harm; perhaps in such a context, though, "hacking" would not be the appropriate term to use to describe the amputations). Does the audience believe the film's continuous depiction of limb-hacking, whether in the context of some deranged psychopath or a conscientious army field surgeon, realizes some property they disapprove of in virtue of its treatment of a morally pertinent matter? Whether they do or not will determine whether they perceive the fiction is in bad taste. If they do consider it to be in bad taste, then whether they believe it has an aesthetic quality that mitigates this bad taste will determine whether in their eyes, at least, it is elevated to good bad taste, rather than bad bad taste. It remains an open question, however, as to how much the recognition of the aesthetic quality of the poor taste depiction mitigates the offense taken (i.e., attenuates the feeling of offense). Certainly, if what Waters has to say is to be believed, it does not negate the possibility of one throwing up. Perhaps one can *feel* offended (because of the property one believes *x* is realizing), just as one can feel nauseous, while also appreciating the aesthetic quality of what is causing the offense or the nausea.

As a final point, recall from chapter 3 the example of people who would regularly faint during readings of Chuck Palahniuk's short story "Guts." Recall, also, that I said that what Palahniuk depicts in his story is not immoral but is horrific. It is therefore possible for audiences to experience a strong aversive reaction to depictions that are (i) neither immoral nor in poor taste (what Palahniuk depicts is not in poor taste because the depiction does not concern the *treatment* of a morally pertinent matter) and (ii) arguably, of high aesthetic quality.

CONCLUSION

As noted at the start of this chapter, in many respects the argument presented here should be viewed as a precursor to the more detailed discussion and argument to follow in chapter 13. In this later chapter, I will have more to say on the issue of poor taste and its status as a suberogatory act when contrasting it with immorality. I began to discuss poor taste, here, however, because matters pertinent to the issue of taste were raised in the previous chapter that I felt needed to be addressed, even if not fully, before moving on.

What I hope to have established through this relatively brief discussion on the matter of taste is that the process by which we arrive at judgments of poor taste is attitudinal, and therefore what pronouncements about poor taste actually mean is that one has a negative attitude toward *x*. What I also hope to have shown (for reasons that will become clear when contrasting poor taste with immorality in chapter 13) is that this attitude arises because of one's belief that *x* realizes some property (O) of which one disapproves: a disapproval borne from one's perception of how *x* is treating a morally pertinent matter (say, in a trivializing or irreverent manner). Before discussing any of this further, however, I would like to move away from an examination of pure fictional immorality and introduce depictions that feature historical figures and/or events. What impact, if any, does the contamination of pure fictional immorality have on the morality of immoral fiction?

NOTES

1. Reprinted by permission from Springer Nature: Springer Nature, *Ethical Theory and Moral Practice* (Young, Garry. An Expressivist Account of the Difference between Poor Taste and Immorality. *Ethic Theory Moral Prac* 22, 465–482 [2019]. https://doi.org/10.1007/s10677-019-09998-2), Copyright © 2019, Springer Nature (2019).

2. To illustrate, as part of a 2012 spoof appeal called "Radi-aid," people from African countries were asked to donate radiators to help the freezing children of Norway. It was devised by Norwegian students as a way to highlight ill-conceived charity campaigns that use stereotypical images from African countries. Out of context, however, the spoof may have been considered "poor taste." In the UK in 2001, the TV show *Brass Eye* ran a spoof current affairs program entitled "Paedogeddon" that tricked celebrities into endorsing what, with hindsight, were clearly bogus campaigns warning of the danger posed by pedophiles. The target of their spoof, and hence amusement, was not the victims of pedophiles nor pedophiles themselves but, rather, the celebrities who were willing to endorse these campaigns with (allegedly) little or no thought about their dubious content.

3. Saying this does not discount the possibility that they understood that others may find the sand sculpture (or whatever *x* represents) offensive.

4. Recall from chapter 2 that what realizes or instantiates O are certain properties ascribed by the creator of the fiction to the fiction (e.g., trivializing a morally pertinent matter).

5. Both examples refer to what might be called, broadly construed, a practical ought (characteristic of Kant's hypothetical imperative) or more specifically a teleological ought in the former case and (possibly?) a prudential ought in the latter (see Chrisman, 2016, for a detailed discussion).

6. One could go further and say that the concept of immorality necessarily contains this obligatory feature—namely, that *x* must not be done—such that, analytically, when stating "*x* is immoral," the further announcement that *x* "ought not to be done (in the obligatory sense discussed)" is redundant. Moreover, even if it could be argued that the same analytic feature is available in the case of "*x* is in poor taste," the nature of the "ought" that necessarily flows from this analysis remains ambiguous.

7. Consider, also, Kenneth Goldsmith's 2015 poem *The Body of Michael Brown*, which is a reworking of the official autopsy report on the death of Michael Brown: an eighteen-year-old black man killed by a police officer in Ferguson, Missouri, in 2014. Attitudes differ over the morality of this poem, given its nature and the fact that it was written by a white man (see, for example, https://overland.org.au/2015/03/when-poetry-is-racist/; accessed 23/01/20). For the poem to be in poor taste, one would need to disapprove of its treatment of a morally pertinent matter. Is the poem exploitative? Does its clinical focus on the *body* of Michael Brown risk losing sight of Michael Brown's humanity? Does it transform Michael into a physical object reminiscent of the way black people were treated as slaves?

Chapter 10

Historical Fiction and Fictional Immorality

INTRODUCTION

In the 1984 film *Amadeus* (directed by Miloš Forman and adapted from Peter Shaffer's screenplay), the chief protagonist, Antonio Salieri, is portrayed as someone who is in awe of the prodigious talent of fellow composer Wolfgang Amadeus Mozart. Yet so consumed with envy and so resentful is Salieri that the self-confessed champion of mediocrity plots to kill him.[1]

Amadeus received eight Academy Awards, including Best Picture, Best Adapted Screenplay, and Best Actor for F. Murray Abraham (who played Salieri). As a work of *fiction*, it is hailed by many as a triumph,[2] although it has been criticized for its historical inaccuracies (Brown, 1992; Keefe, 2009).[3] So how much of *Amadeus* is based on fact?

Salieri and Mozart are genuine historical figures and were musical rivals. Certainly, there is evidence that they met and even socialized on occasion (Borowitz, 1973). After Mozart's death, Salieri remained acquainted with Mozart's family (he taught Mozart's son music, for example, at the behest of Mozart's widow). But is the fictional Salieri's self-confessed mediocrity likewise based on his actual mediocrity, or at least his belief to that effect? Was Salieri actually envious of Mozart, and did he really kill him? There is no evidence, beyond the rumors and speculation that circulated soon after Mozart's death, and resurfaced periodically, that Salieri had anything to do with the death of Mozart, although whether he was envious of him is harder to ascertain. Neither was Salieri considered a mediocre talent during his lifetime. Beethoven, for example, happily referred to himself as "Salieri's pupil" (Borowitz, 1973), so it is unlikely that Salieri thought of himself in this way; while it is true that, over time, performances of his work declined, there are a number of factors that could account for this apart from mediocrity (see Cunningham, 2014).

Historical fiction borrows the lives of historical persons in order to create fiction (Mills, 2000). *Amadeus*, as we have seen, borrows the lives of Mozart and Salieri (among others). In this chapter I examine the morality of historical fiction, particularly in cases of historical figures depicted as engaging in fictional immorality (e.g., Salieri murdering Mozart). To do this, I consider what is involved in depicting fiction from within a historical context and show how much of this is an irrelevance to questions relating to the morality of depictions of fictional immorality. I also revisit the matter of accidental truth in fiction (first raised in chapter 2) and argue that, unlike pure fiction, historical fiction can be accidently true. Importantly, though, knowledge of its accidental truth necessarily changes how we engage with it. Finally, given that the type of historical fiction I have in mind features fictional immorality (i.e., the murder of Mozart by Salieri in *Amadeus*), it is my contention that the same moral concerns that arise from cases of pure fictional immorality also apply to this type of historical fiction. In short, the contamination of pure fiction by cases of historical fiction, so described, does not alter, in any significant way (to be explained), the focus of our moral scrutiny or the extent of our concern.

THE FIDELITY CONSTRAINT REVISITED

If we start from the premise that Salieri did not kill Mozart, then *Amadeus* depicts an act that did not actually happen. The plot essentially takes the form of a counterfactual in which Salieri murders, or at least contributes to the death of (but let us say murders), Mozart (i.e., *if* Salieri murdered Mozart, then this is how it could have happened). Counterfactuals are found in any number of alternate history fictions. Take, for example, the counterfactual in Robert Harris' novel *Fatherland*, in which Germany won World War II, at least in Europe (i.e., *if* Germany won the war, then what follows could have happened). Unlike *Fatherland*, however, in which something that is known not to have happened is depicted as having occurred, *Amadeus* does not depict what is *known* not to have happened but, rather, given the mystery surrounding Mozart's death, what did in fact occur (or an approximation of this) *if* the rumors are to be believed. In contrast, as far as I am aware, there have never been rumors circulating that Germany, despite clear evidence to the contrary, actually won the European theater of World War II, and therefore rumors on which the premise of *Fatherland* is based.

Fatherland intentionally violates the fidelity constraint (recall from chapter 2 that this is a requirement of fiction), not only with regard to the outcome of World War II but also in virtue of all that stems from this in the novel.

Moreover, the violation is known to have occurred by the majority (if not all) of its intended audience. As a member of that audience, I know that Germany lost World War II. To fictively imagine otherwise, as is required to engage with *Fatherland* (i.e., to adopt the authorized work world, to borrow from Walton, 1990), I must suspend this knowledge. In the case of Mozart, however, we cannot claim to *know* that Mozart was not murdered because we do not know what caused his death (see, for example, Davies, 1983; Dupovy-Camet, 2002; Hirschmann, 2001; Karhausen, 2010; Wheater, 1993; Zegers et al., 2009, for a small sample of articles debating this). There is therefore no knowledge to suspend, in this regard, when fictively imagining Salieri's murderous rivalry with Mozart (although, in the absence of knowledge, I may still hold certain *beliefs* that require suspending, such as the belief that Mozart actually died of kidney failure brought on by a streptococcal infection).

What we can claim to know, and therefore the knowledge we need to suspend, when fictively imagining the death of Mozart as depicted in *Amadeus*, is that there is currently no evidence to corroborate the accusation that Mozart was murdered, let alone that Salieri was the murderer, even if this is true (see discussion in the next section). Given this, one could argue that the extent to which *Amadeus* adheres to or violates the fidelity constraint regarding the manner of Mozart's death is not known because the cause of death is not known, and evidence is limited. I concede this point, at least as far as certain facts about Mozart's death are concerned, such as *how* he died. But, because of this, one could also argue that, presently, there *is* a more general fact pertaining to the death of Mozart that we do know: that the exact cause of Mozart's death is unknown. Now, because we can claim to know this fact, in this regard, the portrayal of Mozart's death in *Amadeus* violates the fidelity constraint, because it depicts a specific, albeit unsubstantiated, cause of death—namely, Salieri's foul play—which goes against what we can claim to know about Mozart's death. Of course, as noted earlier, members of the audience may hold different beliefs (with varying degrees of justification) about the death of Mozart. The extent to which an individual's belief about the cause of Mozart's death aligns with the depiction found in *Amadeus* will determine the extent to which they must suspend their belief in order to fictively imagine the truth of what is shown to be the case (again, see discussion in the next section).

In sum, *Amadeus* does not stay faithful to the historical *uncertainty* surrounding Mozart's death but, rather, intentionally violates it by depicting the unlawful killing of Mozart by Salieri. The counterfactual present in *Amadeus* (i.e., *if* Mozart had been murdered then, this is how it could have happened) is specific enough to violate the fidelity constraint, owing to its deviation from the uncertainty surrounding Mozart's death. Because of this, *Amadeus* satisfies important requirements of fiction: the intentional violation of the fidelity constraint (fictive intent) and one's willingness (as an audience) to fictively imagine this violation as a counterfactual condition. The fact that it contains deliberate and significant depictions of historical figures, however, negates its status as pure fiction. In the next section, I discuss how this impacts the status of these depictions as abstract entities and the possibility of events depicted being accidently true.

THE ACCIDENTAL TRUTH OF FICTION REVISITED

Recall from chapter 2 that I presented a qualified rejection of the claim that *pure* fiction could be accidently true. Continuing with the Mozart-Salieri example: Imagine that evidence comes to light that reveals how Mozart died. In this hypothetical scenario, let us allow that he was deliberately killed by Salieri. Suppose that a short time after the cause of Mozart's death is discovered, an old parchment is likewise discovered, dated several centuries before the time of Mozart and Salieri, on which is written a play, set a few hundred years after the author's death, around the time of Mozart and Salieri, about two rival composers who happen to be called Mozart and Salieri, and whose descriptions match exactly the two historical figures yet to be born. Moreover, the latter is depicted as murdering the former. In this newly discovered fiction, "Salieri" acts as an abbreviated descriptor for all that is attributed to the fictional character—such as devout, abstemious, philanthropic, habitant of Vienna, Kapellmeister, composer and music teacher, murderer of Mozart—that also happens to match descriptions we *now* know (in the context of my hypothetical scenario) to be true of the historical figure, Salieri.

Given that the historical Salieri and Mozart had not been born at the time the newly discovered fiction was written, each character constitutes a purely fictional entity in accordance with the necessary and sufficient conditions for pure fiction set out in chapter 2.[4] This means that while what is true *of* each historical figure can be *coincidently* true *of* each fictional character, in virtue of the fact that the abbreviated descriptor and therefore all the descriptions captured by it apply equally to both fictional and actual persons, what is true *about* the historical figures (such as the fact that they *are* the historical Mozart and Salieri and *are* an *historical* murder victim and murderer) cannot be accidently true. Fictional Mozart and Salieri do not represent any historical figures, or anything beyond themselves (and so descriptions *about* them lack truth-aptness; see chapter 2). A point of clarification: fictional Salieri and Mozart could, of course, be described as historical figures within the fiction, or murderer and victim, or even corporeal and biological entities, but even if any or all of this were the case, it would simply be true *of* the attributes the creator of the fiction describes them as having; it is not a truth *about* the fictional entities themselves because, to reiterate, they are not intended to represent anyone outside of themselves, and so cannot possess accidental truths *about* themselves on account of truths about any actual entity.

In the case of *Amadeus*, however, while still in the context of the hypothetical scenario in which the cause of Mozart's death is known, what we (as the audience) now know is that the counterfactual depicted in the film—that Mozart died at the hands of Salieri—coincides with what has recently been revealed to be true, just as the rumors claimed all those years ago. The alignment of Mozart's death in *Amadeus* with the newly discovered facts of the matter could not have been knowledge-driven at the time of *Amadeus*' creation, however, because knowledge of the exact cause of Mozart's death was not available when the screenplay was written (although I accept that the screenplay was no doubt inspired by the rumors of Salieri's role in Mozart's death). Does the alignment of fiction with reality nevertheless make the depiction of Mozart's death in *Amadeus* accidently true? After all, the fictitious entities that appear in *Amadeus* are intended to represent historical figures: to represent something actual, outside of themselves.

The content of *Amadeus*, including *its* relationship to the actual event (death of Mozart), did not change after the discovery of the cause of Mozart's death. What *Amadeus* depicts, before and after the discovery, coincides with actual events *as they relate to the historical figures featured in the fiction*. What changed at the moment of discovery was our knowledge of this relationship. We now *know* that the death of Mozart, as portrayed in *Amadeus*, is an accurate depiction of Mozart's death (something it had always been; we just failed to know it at the time).[5] What the discovery also means is that the depiction no longer violates the fidelity of the historical *uncertainty* surrounding Mozart's death, and therefore the fidelity constraint regarding *how* Mozart died, because this uncertainty has now evaporated.

This last point is important because, according to Kajtár (2017), if the depiction is now known to adhere to, rather than violate, the fidelity constraint, then it is not (no longer) a fiction that invites fictive imagining. Thus, while it is true that *Amadeus* contains the same content now as it did before the cause of Mozart's death was discovered, it is also true that the audience's relationship to this content has changed since the discovery. Given that the audience now knows that what is depicted does not violate the fidelity constraint and is not therefore a counterfactual condition, fictive imagination is not required to engage with the content, even when once (i.e., prior to the discovery) it would have been. It is now known that *Amadeus* contains the dramatization of actual events, and while these are distinct from the actual events they dramatize, engaging with them no longer requires fictive imagination. Put differently, given what the audience understands about the depiction of Mozart's death in *Amadeus*, the content should elicit from the audience belief in the truth of what is depicted rather than its make-believe truth. In answer to the question about the accidental truth of fiction, then, only this time in the context of *historical fiction*: a fiction can be accidently true; it just cannot knowingly be accidently true and remain the same sort of fiction.

What this means for *Amadeus*, at least in the context of the hypothetical scenario we have been discussing, is this: The discovery of the cause of Mozart's death brings with it a transformation, but it is not a transformation in the content per se (i.e., Salieri is still depicted killing Mozart). Nor is it a transformation in the truth status of the content in relation to the actual death of Mozart, which remains the same. Instead, what has changed is our knowledge of Mozart's death and the fact that we now know that *Amadeus* depicts what actually occurred, something that (we now know) was as true before the discovery as it is after it.

The transformation that occurs in the epistemic relationship we have with the content of the fiction affects the depiction's fictional status. After the discovery, *Amadeus* ceases to be a fiction that invites us (requires us, in fact, if we are to engage with the fiction as originally intended) to make-believe the truth of Salieri's role in Mozart's death and, instead, becomes (is transformed into) a dramatization of actual events. The change is not a physical change, of course; rather, it amounts to a change in what it is appropriate for us to believe: in this case, that what is depicted actually happened (Kajtár, 2017).

Returning to the truth of/truth about distinction: What is true *about* the historical figures, Mozart and Salieri, is that they are or, rather, once were corporeal and biological entities *about* which descriptions are truth-apt. What is true about the Salieri and Mozart in *Amadeus* (as well as the composers in my fictitious play who just so happen to be called Mozart and Salieri) is that they are abstract entities *about* which descriptions are not truth-apt. Nevertheless, in *Amadeus*, unlike my fictitious play, what is also true *about* these abstract entities is that they are intended to represent historical figures. This is not true *of* them within the fiction; that is, they are not described in the film as fictional characters intended to represent historical, corporeal, and biological figures but, instead, are ascribed the same features as the historical figures through the creator's descriptions. What is true *of* the fiction is therefore true *of* the actual (at least, in principle, even though, as noted earlier, historical inaccuracies have been reported in the case of *Amadeus*). In contrast to my fictitious play, then, what is true *about* Salieri and Mozart in *Amadeus* is that they are abstract entities constituted to represent a historical murderer and his victim. Now, while this *particular* representation (i.e., murderer or victim) may have been intended as a fiction by its creator (in accordance with a violation of the fidelity constraint), within the wider historical context portrayed, it just so happens (on account of the introduction of my hypothetical scenario) that what is true *about* these abstract entities in terms of what they represent is accidental true, rather than true by design (i.e., the historical persons they represent as murderer and victim were, in reality, just that). Once the audience learns of this, of course, as already noted, the epistemic relationship between audience and fiction is transformed, as is the nature of the fiction itself. As a member of the audience engaged with the dramatization of historical events, I am required to believe that what is depicted is an accurate representation of historical fact, rather than fictively imagine, and therefore make-believe, the truth of what I engage with. As such, I am no longer required to willingly disregard my belief that Mozart died from self-inflicted mercury poisoning while being treated for syphilis (for example) in order to make-believe that he was murdered by Salieri. Instead, I am required to believe that he was murdered in the manner depicted in *Amadeus*, as well as everything else depicted therein (let us allow) as a matter of historical record.

ON THE MATTER OF ACCURACY

The dramatization of historical events, even when portrayed accurately—that is, A is depicted murdering B and did in fact murder B, historically—are inevitably going to be presented in a way that is colored by a particular interpretation of the event, and so include what Lukács (1983) calls *necessary anachronisms*. Some of these will be self-conscious, and therefore intended, while others will be unintended. To illustrate: in the 1986 UK comedy *Black Adder II* (written by Richard Curtis and Ben Elton, and set in Tudor England), the protagonist, Lord Black Adder, sometimes refers to his sidekick, Lord Percy, as a prat: "You really are a prat, aren't you Percy."[6] To have someone utter this term in a comedy set during the reign of Queen Elizabeth I (circa sixteenth century CE) is anachronistic and would fail what Demos (2005) calls the "ring true test" (p. 332). It is intentionally included, however (I surmise), to convey Black Adder's displeasure at Lord Percy to a contemporary audience in a manner they would find amusing, not least *because* of the anachronistic language spoken. In contrast, the sexism evident in the previously mentioned 1960s series *Star Trek* (see chapter 4) reflects a not uncommon attitude men had toward women at that time and was likely unintentionally captured within the fiction. To a modern audience, it appears anachronistic and (as noted previously) is, at times, uncomfortable to watch.

Given the unavoidable influence of one's own perspective on the interpretation of historical events (which produces the aforementioned anachronisms), suppose, a new dramatization of the life of Australian bushranger and outlaw Ned Kelly is commissioned (Ned Kelly was convicted of the murders of three police officers and hanged in 1880). No doubt, there are different ways to interpret the actions of Ned and the rest of the Kelly gang. One may adhere to the fidelity constraint with regard to physical events when dramatizing his life (i.e., correctly depicting the three murders Kelly was convicted of carrying out, as well as other deeds), but what does it mean to adhere to the fidelity constraint regarding the *interpretation* of the killing of the police officers and the events that led to this? In other words, what motivated Kelly to commit these violent acts? Even today, Ned Kelly is a hero to some and a villain to others, just as he was in his own lifetime.[7]

Harrold (2003) refers to historical fiction of the Ned Kelly variety as "accuracy aspiring" (i.e., one aspires to provide an accurate rendering of the facts as well as, say, the psychology of the protagonist) and contrasts this aspiration with other "broadly realistic" fiction that does not prioritize accuracy.[8] As he states:

> [Accuracy aspiring] works are [to be] distinguished from broadly realistic works such as Shakespeare's *Henry V,* because they purport not only to portray psychologically realistic characters and plausible actions and events, but to do much more: to accurately report major historical events and social/cultural facts when they are relevant, and sometimes to describe the thoughts and actions of particular characters in ways that are not only psychologically plausible, but that also closely approximate their *actual* actions or thoughts (or the actual actions or thoughts of other persons of the type being imagined). (Harold, 2003, p. 247)

We must be cautious when evaluating the veracity or authenticity of such accuracy-aspiring fictions, however. Not because we have reason to doubt the accuracy of the depictions, at least as far as the physical aspects of the historical events are concerned (i.e., who did what to whom, when, and where) but, rather, because history, in its fullness, contains many features that are less exposed and therefore less accessible to public scrutiny, such as the thoughts, feelings, and motivations of the key players. Does Arthur Penn's 1967 film *Bonnie and Clyde* accurately portray the psychological make-up of Bonnie Parker, for example? The way Parker is portrayed in Penn's film is certainly in stark contrast to what is, for the most part, implied within John Lee Hancock's 2019 dramatization *The Highwaymen*: that she was a cold-blooded killer. Which, if any, of the two dramatizations more accurately depicts the actual Bonnie Parker? But, equally, if one's goal is to create a *fiction* within a historical context, what does it matter? After all, and as a way of justifying the fictional component of historical fiction:

> A simply true account of the world based on what the documentary record permits one to talk about . . . can provide knowledge of only a very small portion of what "reality" consists of. . . . [T]he rest of the real, after we have said what we can assert to be true about it, would not be everything and anything we could

> imagine about it. The real would consist of everything that can be truthfully said about its actuality plus everything that can be truthfully said about what it could possibly be. (White, 2005, p. 147)

In addition, and with reference to Primo Levi's writings on his experiences as an inmate of Auschwitz concentration camp during World War II, White has the following to say about the use of imagination:

> The significance of Levi's book lies less in any new "truthful" information he gives about the camps than in the artistry (by which I mean literary, poetic and rhetorical devices) he employs in order to conjure up a compelling image of a cosmos utterly horrifying and at the same time horrifyingly present as a possibility for everyone of our time. Do I mean to imply that Levi's account of his year in Auschwitz is a fiction in the sense of being a pure invention? Of course not. . . . Levi manages to demonstrate to his readers the difference between a merely truthful account of an event, of the kind provided by most survivor-witnesses, and an artistic treatment of a real event in his past which transcends the truth-reality distinction. (*ibid*., p. 149)

For White, in addition to the historical accuracy of Levi's work, he remains faithful to the horror *experienced* by those having to endure life in a Nazi concentration camps. This is not something that can be conveyed in its fullness, I contend, simply as a list of events but, instead, requires skilled imagination on the part of the writer/creator to capture what the facts alone cannot; and is no less true for that.[9] I am not suggesting that a skilled historian is incapable of bringing history to life in this way, only that "[one can acquire] a love for the atmosphere of the past through the imagination of a great storyteller" (Merritt, 2014, p. 1)

What is also true is that historical fiction offers a perspective on history, irrespective of whether it aspires to be accurate or is less stringent in this regard; and while this may pique the interest of its reader to pursue other nonfictional accounts (Slotkin, 2005), it also seeks to entertain (McGarry & Harriman White, 1963). To illustrate: Mel Gibson's 1995 film *Braveheart* won five Academy Awards, including best film (tentatively, I will take this, and my own recollection of the film, as evidence that it is highly entertaining). After its release, it also raised interest in the profile of the film's protagonist and genuine Scottish historical figure and hero William Wallace (1270–1305), and well as interest in Scottish history and tourism more generally.[10] The film is, however, full of historical inaccuracies. An inaccuracy of less concern, perhaps, is the fact that the French love interest in the film—Princess Isabelle—never actually met Wallace and was, in fact, only thirteen years old when he died (not in her late twenties, as portrayed in the film). Nor, perhaps, is it a major concern that the Scots warriors were unlikely to have worn blue face paint during this period, nor kilts. Of greater concern, historically, and for those proud of their Scottish past, is the fact that Robert the Bruce (who features in the film) was actually the person referred to as Braveheart and not William Wallace, and also that Bruce (who is also regarded as a true Scottish hero) did not betray Wallace at the battle of Falkirk, despite being seen to do this in the film.

When engaging with historical fiction like *Braveheart*, as with any fiction, we are invited to make-believe what is being depicted within the context of the overall story. Braveheart is meant to be a fiction, after all, even though it is based on *some* historical facts. But within the overall make-believe, what we are invited or otherwise come to *believe* about the historical account on which the fiction is based may be of moral significance, as champions of Robert the Bruce would no doubt testify regarding the manner of his portrayal in the film. A further example used by Cooke (2014) makes essentially the same point, only with wider implications:[11]

> Director Kathryn Bigelow and writer Mark Boal make it clear that their film *Zero Dark Thirty* (2012) was created and presented with values of sincerity and accuracy. The film opens with the words "Based on Firsthand Accounts of Actual Events," and although it is clearly a fiction (for example, its central character is a composite of several actual individuals), it signals that its makers are sincere and have striven for accuracy. This is precisely why the film, despite being a fiction, has become so controversial—it suggests that torture was essential in collecting the intelligence that ultimately led to Osama Bin Laden, when every reliable person involved in that hunt averred that any such measures led away from their target. (p. 321)

Cooke's point is that, in *Zero Dark Thirty*, there is a risk that instead of make-believing that the torture led to the gathering of intelligence on the whereabouts of Osama Bin Laden, viewers may come to *believe* that this is how it was actually acquired. There is also the risk that the use of torture will not only be perceived as a fixed norm (Booth, 1988; see chapter 4) and therefore as something it is acceptable to do under certain circumstances (see, for example, Brown, 2007; Dershowitz, 2004; Gert, 1969; Shue, 1978; Sussman, 2005, for discussion on the (alleged) legitimate use of torture) but that, in the case of *Zero Dark Thirty*, it was an *effective* means of obtaining information about Bin Laden and his associates. Such a perception would be in stark contrast to what actually occurred when the numerous detainees were tortured. Belief in the effectiveness of torture (for gathering intelligence), if elicited from the film's portrayal of torture, also runs contrary to considered opinion on the topic (see, for example, Bell, 2008; Bufacchi & Arrigo, 2006;

Lowth, 2017; O'Mara, 2015). As Langbein (2004) concludes: what an extensive study of history has taught us is that it has not been possible to make coercion compatible with truth.

While there is much more that can be discussed on the subject of accuracy-aspiring dramatizations of historical events (see de Groot, 2010; Gethering, 2012; Harold, 2003; Slotkin, 2005, by way of a small sample), such fiction is not of interest to my continuing examination of fictional immorality. This is because, if the dramatization is accurate, as accuracy-aspiring dramatization tries to be, then any immorality depicted is not going to be of a *fictional* kind. (I say this while acknowledging the aforementioned concern over interpretation and how the interpretation given to a particular depiction may itself be regarded as a work of fiction by some and appropriate by others.) Because of this, I am interested in historical fiction that intentionally violates the fidelity constraint by depicting historical figures engaged in *fictional* immorality, as is the case with Salieri in *Amadeus* (outside of my hypothetical scenario) and, as one among many other examples, Charles Lindbergh in Philip Roth's 2004 novel *The Plot Against America*.[12] Are there any moral concerns associated with historical fiction of this kind, beyond those already raised in the preceding chapters when discussing pure fictional immorality?

ON THE MORALITY OF HISTORICAL FICTION

At the start of the chapter, I noted how historical fiction borrows the lives of historical persons in order to create its particular brand of fiction. Given that historical fiction depicts historical figures and/or events, and given that the type of historical fiction I have in mind necessarily violates the fidelity constraint with regard to depictions of fictional immorality, is the act of borrowing historical figures/events for the purpose of depicting fictional immorality morally problematic?

For Slotkin (2005), historical fiction is a useful means of exploring alternate possibilities or *what ifs*. What if Salieri had murdered Mozart? *Amadeus* gives us insight into what this might have involved. Similarly, *what if* Charles Lindbergh had been elected to the office of president of the United States in 1940? Given Lindbergh's documented political views, in which direction would he have taken America during the most violent and bloody period in European history? *The Plot Against America* explores one possibility.

In many respects, alternate histories are nothing more than elaborate thought experiments, and, as an intellectual pursuit—*qua* the *examination* or even simply the delineation of an alternate history—they are amoral. In its role as educator, the fiction informs our understanding of history, and, in accordance with this role, *even as a work of fiction*, one might expect—possibly require—that it respects and strives for authenticity (Groce & Groce, 2005). Yet historical fiction is also meant to entertain and perhaps even maintain a high aesthetic quality (McGarry & Harriman White, 1963); and while authenticity and entertainment/aesthetic quality are not mutually exclusive, neither is the former synonymous with the latter. A potential for tension arises, therefore, between these independent elements, as a critical comment by Hazel Rochman illustrates, when reviewing Katherine Lasky's historical fiction *Beyond the Burning Time*: "The history overwhelms the fiction, although both are compelling" (cited in Brown & St Clair, 2006, p. 42). Consequently, what if it transpires that the authentic is not as entertaining/aesthetically pleasing as the inauthentic? Should we sacrifice one for the sake of the other?

If one wishes to prioritize the educational component, then it follows that authenticity should not be sacrificed for the sake of entertainment or a higher aesthetic appeal, and by "authenticity," I do not mean authenticity toward the historical facts alone but also authenticity of *experience* (recall the discussion on Primo Levi's writings about Auschwitz). Where one seeks, first and foremost, to entertain, however, and/or attain a higher aesthetic standard and, in doing so, opts, to a greater or lesser degree, to sacrifice authenticity in one or more of its guises, then is there a moral price to pay for this sacrifice?

In response, consider the following anonymous blog comment:

> [What] bothers me is when an author takes a historical person or event and intentionally represents it differently. . . in order to write some other story or convey some other point—using the established cultural meaning of a real person or event to lend color to his fiction. . . . [T]he danger [is that] it runs the risk of obscuring . . . the real human events and dramas that people experienced. And when we don't know what really happened. . ., in a certain sense, we no longer know who we are or how we got here.[13]

The concern of this anonymous blogger is that historical fiction is not only parasitic on historical figures/events but also that it writes a different "story" to the one told by historians. By misrepresenting history, historical fiction risks making it harder for the real story to be known. A similar view is expressed by another blogger (abramdemski): "Fiction is not a lie, but it is a variety of untruth. It absorbs time and energy which could be spent on fact."[14]

Of course, historians often disagree over how to interpret historical figures and events, and revisionist accounts appear periodically that challenge established views in the hope of achieving a heightened state of *verisimilitude*. In this latter regard, accuracy-striving historical fiction and revisionist history have a common goal. But the historical fiction I seek to examine does not strive for verisimilitude. As the anonymous blogger points out, it *intentionally* misrepresents historical facts, at least as they are currently understood.

No doubt a large part of the interest, intrigue, and entertainment value of historical fictions like *Amadeus* and *The Plot Against America* stems from the fact that they feature genuine historical figures. The fiction therefore trades on the lives of people who have already secured a place in our history. Yet, in return, what does it benefit Salieri or Lindbergh to be depicted as the chief protagonist in a counterfactual (and therefore fictional) condition in which they engage in immorality? Leaving aside the fact that both were deceased long before the publication of the respective historical fictions that feature them, Salieri was well known in his own lifetime: a celebrity by today's standards. Lindbergh, for his part, was also famous for making the first solo transatlantic flight (but also because of the tragic death of his son, who was kidnapped and murdered). He also championed the isolationist movement in America that began during the 1930s and was accused of being anti-Semitic and a Nazi sympathizer, although the latter accusation was more contentious than the former (Olson, 2013).

Neither Salieri nor Lindbergh lacked celebrity status. So, even if they had been alive at the time of each fiction's creation, neither person, it would seem, needed the added publicity to elevate their status nor, I imagine, would they have craved the notoriety. Because of this, is the borrowing of an historical figure's life (or an event) in order to misrepresent it for the purpose of entertainment, even if this is not the sole purpose of the fiction, exploitative?

It is not unreasonable to conjecture that each of the historical figures we are currently discussing would be (relatively) unknown to a modern audience. Historical fiction of this kind (involving counterfactuals) could therefore prove to be a way for an unfamiliar audience to learn more about these historical figures, should the audience choose to move beyond the misrepresentation constitutive of the counterfactual by seeking history texts that are nonfictional and seek not to misrepresent. Prior to, or in the absence of, further research by the audience, however, have Salieri and Lindbergh been harmed by the manner of the fiction's depiction of them? Is the depiction harmful irrespective of the audiences' knowledge of history?

There is no evidence (that I am aware of) to indicate that historical fiction of the kind I am discussing has ever endorsed the counterfactual condition it depicts. *Amadeus* presents its audience with representations of historical figures engaged in a fiction without explicitly accusing the actual (and long dead) Salieri of Mozart's death. The film does not come with an expectation that the audience will proclaim *salve veritate* after watching it, although this fact may not prevent some from misinterpreting what is depicted, anyway (i.e., as the dramatization of an historical event). The lack of endorsement is even more striking with *The Plot Against America*, given that one of the central premises of the fiction—that Charles Lindbergh became the president of the United States—is known to be historically false. It is therefore an open question as to whether such historical fiction hampers one's knowledge of the favored historical account. (Recall, this was a complaint of the anonymous blogger.) But even if it were the case, would such hampering be evidence of (i) a failing on the part of the *fiction* as opposed to, say, the audience and (ii), if it is a failing, is it a *moral* failing?

It is not a failing on the part of the fiction because the fiction does not try to pass itself off as a factual account of history and, for this reason, should not be obliged to present authentic depictions. Hampering, if it is to be found, should be taken as evidence of the audience's willingness to content themselves with a fiction-based account of history, or at least the historical figure(s) depicted. But even if we shared the failing between the audience and the fiction—by declaring that the fiction should have been clearer about how far it deviates from historical records and the audience should have made more of an effort to pursue these records and therefore learn more about the actual lives of the historical figures depicted—is it a *moral* failing by either party? No, it is not.

It is not a moral failing if there is no intention by creator or audience to deceive, mislead (I do not include disclosed misrepresentation, here), besmirch, or otherwise disrespect the actual person featured in the fiction. That said, where the intention, and therefore the meaning of the fiction's message in relation to historical accuracy, is ambiguous, then the fiction is vulnerable to a charge of moral insensitivity, as discussed previously; see, also, below.

Objections could still be raised against the fiction, even if one accepts that what is portrayed is not intended to be historically accurate. I will call the first of these objections "No smoke without fire" and the second "Why would you even *think* that?"

In the case of *Amadeus*, the "No smoke without fire" objection would look something like this. One could argue that Salieri's reputation is sullied, even if one accepts that there is no evidence to support the truth of what is depicted in the film, because what is depicted helps articulate and animate a possible explanation of a death whose cause is (currently) unknown. After all, there were rumors circulating at the time about Salieri's role in the death and *there is no smoke without fire*. Alternatively, or even concurrently, with the "How could you even *think* that?" objection, one could object to the very idea of what is being depicted, even as a fiction, because it is preposterous. Of course, one could retort that the depiction merely illustrates the contingent nature of history, and there is no logical or metaphysical reason to deny the possibility depicted. It is not preposterous in that sense. The would-be objector may not, however, be objecting to the dispassionate contemplation of an alternate, metaphysically possible, outcome—in the guise of, say, a purely academic counterfactual thought experiment—but to a perceived insinuation that stems from a false characterization of the actual historical figure in virtue of the fictional immorality they are depicted to have been engaged in. And given that the historical figure is now deceased, if the insinuation is believed, what is left to be harmed is the deceased's reputation and, perhaps by association, the family.[15]

We have already discussed the mischaracterization of Salieri in *Amadeus*. In the case of Lindbergh, one may object to the idea that any documented evidence of anti-Semitism or comment that could be construed as sympathetic to the Nazi party is sufficient justification for the, albeit fictitious, portrayal of Lindbergh by Roth in *The Plot Against America*. In effect, one might claim that despite the documented views of Lindbergh, it is nevertheless preposterous to think that he could be capable of what is being suggested, even hypothetically. In addition to the "How could you even *think* that" objection, then, the "No spoke without fire" objection is also germane, insofar as one could argue that what the fiction is doing, even as a recognized fiction, is extrapolating from Lindbergh's known political views about race (for example), to create an alternate version of history that is not implausible, thereby tainting Lindbergh's reputation (i.e., this is what could have happened if someone like Lindberg had become president).

With each example of historical fiction, one might add the further objection that any suggestion of impropriety through the fiction is not done for the purpose of enlightenment, and therefore to attain a heightened state of verisimilitude, but *solely* (which is an important word to include) for entertainment: entertainment at the expense of reputation.

Recall from chapter 7 that enjoying that which is immoral is an immoral thing to do. For such a charge to gain traction, then, it needs to be established that what one is enjoying is immoral. How this is determined in the case of historical fiction is, however, no different to how one would judge the morality of fictional-x_i: namely, by an examination of content, meaning, and harm.

As with fictional-x_i, I see nothing in the *content* of historical fiction, even where it depicts fictional immorality, to merit a judgment of immorality. Likewise, when examining the *meaning* of the historical fiction, following the discussion in chapter 4, one can endorse, in principle, IWV (namely, where a fiction (in this case, historical fiction), taken as a whole, endorses an immoral worldview, then any depiction of immorality from within the narrative, congruent with that worldview, is immoral) while at the same time accepting that difficulties arise when trying to implement it. Given this difficulty, the best we can claim is $\mathrm{WV}_{\mathrm{insensitive}}$, which states: where a reasonable interpretation of a fiction (again, in this case, historical fiction), taken as a whole, cannot establish whether the fiction is merely depicting rather than endorsing an immoral worldview, then any depiction of immorality from within the narrative congruent with that worldview is, minimally, morally insensitive.

The discussion on harm in Chapters 5 and 6 is equally applicable to historical fictional, but with an added aspect. Given that we are now dealing with depictions of historical figures, harm to the person's reputation also needs to be considered. Both the "No spoke without fire" and "How could you even *think* that" objections are based on the view that the fiction is seeking to elicit from its audience belief rather than make-belief that, *if* the contingencies of history had been different, the depicted counterfactual is *plausible* rather than simply metaphysically possible. Where the audience does not buy into the belief about plausibility, then the historical figure's reputation will be unaffected. This does not excuse, morally, the creator of the fiction, however, if their intention was to elicit the belief in its plausibility, or even entertain its plausibility for the sole purpose of entertainment. If so, one may be guilty of violating $WV_{\text{insensitive}}$.

CONCLUSION

Historical fiction involving intentional depictions of fictional immorality does not present a special case. The contamination of pure fictional immorality, at least of the type fitting our interest, has not produced moral concerns beyond those already discussed in previous chapters. The only slight exception to this claim is perhaps one of concern over the actual person's reputation (posthumous or otherwise) and other harms that may befall them on account of damage to their reputation (more so, if they are still alive). Such concern is dependent on the extent to which a particular narrative involving fictional immorality is being endorsed rather than presented as make-believe or where the possibility of endorsement is ambiguous. Each of these possibilities has, however, been examined more generally in previous chapters and can be adequately applied here (e.g., the conditions for moral insensitivity).

Historical fiction is not the only fiction that is capable of depicting actual persons engaged in fictional immorality, of course, and therefore not the only fiction that may potentially damage someone's reputation or cause harm in other ways, even where what is depicted is not necessarily an example of fictional *immorality*. In the next chapter, I consider the moral implications of a much more recent and original form of fiction that likewise depicts actual people: *deepfake* pornography.

NOTES

1. See also Pushkin's 1830 poem, "Mozart and Salieri," which depicts Salieri in a similar light (Pushkin, 1933).

2. For example: https://www.theguardian.com/film/2011/dec/20/my-favourite-film-amadeus

https://www.empireonline.com/movies/reviews/amadeus-review/ (accessed November 16, 2019).

3. See, also, news articles disputing the depiction of Salieri in Amadeus.

https://www.reuters.com/article/uk-austria-mozart-salieri/mozart-museum-seeks-to-debunk-evil-salieri-poison-myth-idUKBREA0T18920140130

https://www.newyorker.com/magazine/2019/06/03/antonio-salieris-revenge

http://www.bbc.com/culture/story/20150224-what-amadeus-gets-wrong (all accessed November 17, 2019).

4. I am assuming that the fictional Salieri and Mozart were not based on actual persons known to the author of the play.

5. We could not say the same thing about the (fictitious) recently discovered play about two composers' names Mozart and Salieri, however, because the play does not depict *historical figures*.

6. Taken from the episode entitled "Head."

7. For an actual example of historical fiction that offers a controversial interpretation of the actions/motives of historical figures, see Kenneth Roberts' 1933 novel, *Rabble in Arms* (Camden, Maine: Down East Books) about the American War of Independence.

8. Of course, even the most well-intentioned, accuracy-aspiring historical fiction may contain inaccuracies. In Geoffrey Trease's *Mist of Athelney*, for example, which is set in England during the ninth century, characters are depicted eating a meal of rabbit stew. This is unlikely to have occurred, however, as rabbits did not populate the wild in England until around the twelfth century, although the Romans are believed to have brought some over with them centuries earlier (example taken from Brown & St Clair, 2006, p. 43).

9. See, for example, Primo Levi's 1959 book, *If This Is a Man* (Orion press).

10. https://www.bbc.com/news/uk-scotland-tayside-central-34127673 (accessed November 25, 2019).

11. See also von Tunzelmann (2013).

12. In Roth's alternate history novel, Lindbergh defeats Franklin D. Roosevelt in 1940 to become the president of the United States. He then proceeds to sign nonintervention pacts with Germany and Japan and oversees the rise of anti-Semitism in America.

13. http://darwincatholic.blogspot.com/2013/05/the-ethics-of-historical-fiction.html (November 26, 2019).

14. https://www.lesswrong.com/posts/RS3C6BAaaaHH9DmDi/fiction-considered-harmful (accessed November 28, 2019).

15. I intend to limit my discussion on posthumous harm to the deceased's reputation and, in doing so, freely acknowledge that I am side-stepping the issue of whether a dead person can be harmed in other ways. For further discussion on harming those who are deceased, see Aristotle (1999), Feinberg (1984), and Pitcher (1984).

Chapter 11

A New Kind of Fiction

INTRODUCTION

In this chapter I continue my examination of fictional immorality involving depictions of actual persons. The focus of this chapter is *deepfake pornography*. At the outset, I would like to restate my position on pornography: it is not my intention to examine the morality of pornography broadly construed. I have discussed child pornography and its relationship to virtual and other fictional depictions of (fictional) pedophilia in chapter 3. Here, I intend to discuss a new form of pornographic fiction. Before doing so, however, it is important to note that this new form of pornography, as with pornography more generally, does not necessarily depict fictional *immorality*. Nevertheless, it is simple enough to imagine cases in which what is depicted is a fictional example of immorality: an adulterous affair or nonconsensual intercourse. That said, it is my contention that the discussion points and conclusions drawn in this chapter are applicable to deepfake pornography regardless of whether it depicts fictional immorality or not.

DEEPFAKES: A NEW KIND OF FICTION

Advances in digital technology (e.g., FakeApp) have enabled the effective manipulation of facial imagery; and the ease with which this technology can be accessed and used has, among other things, facilitated the manufacture of *face-swapped porn*, also known as *morph porn* or deepfake pornography (Güera & Delp, 2018).

To create deepfake pornography (hereafter, deepfake$_p$), a deep learning algorithm is employed to construct a novel digital image of a face—often a celebrity's and often female—from numerous publicly available (or otherwise accessible) images, known as a "faceset" (Harris, 2019).[1] The artificially created image is then mapped onto the body of a person starring in a pornographic video. By replacing the original face with the celebrity's, it appears (often quite convincingly) that the celebrity is the one performing the sexual act.

Deepfake manipulations do not always involve pornographic imagery, it must be said; but, according to a report published by *Deep Trace* (a cybersecurity company based in The Netherlands[2]), of the 14,678 deepfake videos found to be available online during a seven-month period between 2018 and 2019 (which was almost double the figure previously calculated), not only were 96 percent found to be pornographic but, importantly, they were judged to have been created without consent (see also Khalid, 2019). Moreover, while most deepfakes$_p$ "feature" celebrities who, as noted, are usually women (Lee, 2018), they have also been produced for the purpose of "revenge porn" (Morris, 2018; Robertson, 2019)—albeit in a fraudulent form—that, again, mostly targets women.[3] Calls have therefore been made for the classification of deepfakes$_p$ as a type of nonconsensual pornography (Delfino, 2019) and, subsequently, as a form of image-based sexual abuse (Henry et al., 2018). In keeping with this view, Melville (2019) refers to deepfakes$_p$ as insidious.

According to Delfino (2019), deepfakes$_p$ violate the expectation that sexual activity should be founded on consent. This claim needs to be examined, however. Let us assume that, in the original version of the pornographic video (from which the deepfake$_p$ is derived), consent was obtained from those who featured in the video. (Given the generally professional nature of the pornographic industry, this is not an unreasonable assumption to make.) They consented to engage in sexual activity and have it recorded and viewed by others. What is far less likely, however, is that they gave their consent, even tacitly, to any use of their image in a future deepfake$_p$ version of the original video, and subsequently to the distribution and viewing of it by others.[4] But to admit this is not to denounce deepfakes$_p$ for violating the expectation that sexual activity is founded on consent. The sexual activity that occurs in deepfakes$_p$ (in the absence of evidence to the contrary) is sexual activity founded on consent—consent that was given at the time of the original recording and therefore at the time of the original sexual activity.

Deepfakes$_p$ do not show a celebrity (for example) actually engaging in sexual activity; neither do they show the same celebrity actually engaging in what *appears to be*, but is not, sexual activity. Instead, they give the *appearance* of showing the celebrity engaging in what is, in fact, sexual activity. In other words, they depict someone who was not part of the original sex act taking part in it. And while it is unlikely that the celebrity, whose face has been substituted for the person's in the original footage, would have consented to the manipulation, as they did not participate in the sexual act recorded, their consent *for this* is not required.

Deepfakes$_p$ are often a convincing deceit, but whatever moral concerns arise on account of this deceit, violating the expectation that sexual activity is based on consent is not among them. Could it be, then, that deepfakes$_p$ violate the expectation that *appearing* to show someone engaged in actual sexual activity, in a manner intended to deceive, is based on consent? The first thing to note about this claim, of course, is that if one were to consent to be involved in an intentional deception, in the knowledge that it was a deception, then one would be morally culpable for the deception. Leaving this matter aside, at least for now (see the discussion on fakery in the section "The Intent to Misrepresent in the Absence or Presence of Disclosure"), the fact remains that deepfakes$_p$ are designed, at the very least, to *misrepresent* the sexual activity of those targeted.[5] It is therefore my contention that *if* appearing to show someone engaged in sexual activity—intentionally misrepresenting them, irrespective of the intention to deceive—is done in a way that is pornographic, then the act of doing this, including the means by which it is achieved, should require consent, even if one's intent is to produce a fiction with artistic aspirations. In other words, even if one's purpose is to create an aesthetically meritorious artifact, given what is involved in its creation (*qua* a deepfake$_p$), one's reach, no matter how ambitious, does not mitigate the need for consent (see, also, the section "Deepfakes$_p$ as Indicative of a Lack of Respect" on the issue of disrespect).

DEEPFAKES$_p$ AS NONCONSENSUAL PORNOGRAPHY

Recall from chapter 3 that an object acquires the ontological status of *pornography* if (a) the object is put to pornographic use and (b) it is reasonable to believe that the object will be used as pornography, in accordance with point (a), by most of the audience for which it was produced (Rea, 2001). Footage from which the deepfake is derived typically contains depictions of sexual activity that is pornographic. It is on account of this that sexual content in the deepfake$_p$, although now presented in a novel form featuring the target of manipulation (the celebrity or whomever), is classified as pornography. Importantly, then, as an independent creation, deepfakes$_p$ are capable of satisfying Rea's conditions for pornography (recall that the same can be said about virtual pedophilia, even though all of the images of children are computer-generated; see chapter 3).

S, of course, may consent to engage in a sexual act and yet not consent to its recording and use as pornography. Or may consent to its recording for personal use, *only*, and so not to its wider distribution as pornography (cases of revenge porn are applicable, here). S's consent for *x* to be used as pornography is therefore independent of S's consent to engage in whatever sexual activity forms the basis for the pornography. Consent in the former case is required in addition to consent in the latter case. Moreover, the fact that pornography featuring S typically requires S to engage in sexual activity is a contingent fact about pornography, not a necessary one, as deepfakes$_p$ demonstrate. After all, in deepfake$_p$, S does not engage in any sexual activity herself.

In a deepfake$_p$ featuring the manipulated digital image of S, which satisfies Rea's (2001) conditions for pornography, *if* consent for the deepfake$_p$ as an independent pornographic creation is not obtained, then it follows that it should be described as *nonconsensual pornography*. The first morally worrisome feature of deepfakes$_p$ is therefore:

1. The lack of consent obtained from the target of manipulation (i.e., the celebrity/noncelebrity) to use their image in the creation of *pornography* (hence, the charge of nonconsensual pornography)

THE INTENT TO MISREPRESENT IN THE ABSENCE OR PRESENCE OF DISCLOSURE

Deepfakes$_p$ invite moral scrutiny for reasons other than their lack of consent, however. A further reason is:

2. The intent to misrepresent in the absence of a disclosure of intent

As already noted, the sexual act depicted in a deepfake$_p$ video is not fictional. In the case of hardcore pornography, sexual activity takes place, and this remains the case in the deepfake$_p$ version. Instead, what is "fictional," insofar as it is created without seeking to adhere to truth conditions, is that it is *this* celebrity (for example) who is engaged in the sexual activity. A point of clarification: in the case of deepfakes$_p$ (at least the ones we are discussing at the moment; see below for alternate versions), given that the intention on the part of the creator is to deceive—to pass something off as genuine when it is not—what has been created does not satisfy the definition of fiction I have been using throughout this book, hence my use of the term "fictional" (in inverted commas) above. In effect, what has been created is better described as fake, rather than fiction. It is genuine *pornography* with fake content. Where the fake is intentionally advertised as genuine, then it should suffer the same moral recriminations as any other fake that sets out to deceive (e.g., fake artwork, money, antiques). Moreover, where one consents to the fakery, as noted earlier, one is morally culpable. This being the case, it seems peculiar to *require* (as was stated above) that the target of manipulation consent to the pornographic misrepresentation, *irrespective* of the intention to deceive. After all, where the intent is to deceive—to pass the deepfake$_p$ off as genuine—would it be better for consent not to be sought so that the celebrity (or whomever) cannot be incriminated should consent be given, as seems to be the requirement?

In response, I would say that the intent to misrepresent and the intent to deceive are distinct intentions (after all, one can intend to misrepresent without deception). Where one intends to misrepresent S in the manner discussed, one should obtain S's consent, and this should be a requirement irrespective of any further intention, such as the intent to deceive: that is, to pass the deepfake$_p$ off as genuine.

Of course, where the intent is fakery, the extent to which someone *would* (irrespective of whether they should) seek S's consent to misrepresent their sexual activity, or seek this but not seek their consent to market the deepfake as genuine, is uncertain. I feel either is unlikely because there seems to be little, if any, advantage in doing so; but, in the absence of evidence, I am reluctant to assert this. Certainly, any denial that the deepfake is genuine could serve as a useful marketing tool: "They would say it's a fake, wouldn't they." Nevertheless, as a point of principle, irrespective of what else one intends, consent should be obtained *if* one intends to misrepresent the sexual activity of someone in a manner that is pornographic because the creation of any form of pornography featuring actual persons, irrespective of whether they are actually engaging in sexual activity, should require consent.

Still on the issue of disclosure, suppose that a much more transparent deepfake$_p$ is created: one that discloses the misrepresentation. In traditional deepfake$_p$ style, a celebrity's face is substituted for the star of a pornographic video—only, this time, it is presented *as* a deepfake$_p$: that is, as a *fiction*. Let us allow, for example, that a caption is always present in the bottom corner of the video disclosing the manipulation. No one who watches the deepfake$_p$ (and who reads English) is left in any doubt that it is not really the celebrity having sex with the pool attendant.

It is still possible for a "transparent" deepfake$_p$ to satisfy Rea's conditions for pornography, meaning it could still be produced and enjoyed as pornography. Being transparent does not negate this possibility. The matter of consent has therefore not changed, as far as the target of manipulation is concerned (for ease of argument, let us allow that all those involved in the original pornographic video have given their consent). Consequently, even when marketed as such, the deepfake$_p$ remains a form of nonconsensual pornography. Objection 1 continues to apply, even if objection 2 does not.

DEEPFAKE$_P$'S POTENTIAL FOR HARM

In addition to the issue of consent, but irrespective of disclosure and the intent to deceive, a further reason to be morally concerned is:

3. The deepfake$_p$'s potential (*qua* nonconsensual pornography) to harm the target of manipulation (thereby making it a form of image-based sexual abuse)

Regardless of the transparency of the deepfake$_p$—that is, even if it is marketed as a deepfake$_p$ and therefore a fiction (in keeping with how I have been using the term)—the target of manipulation (the celebrity/noncelebrity) could still suffer harm as a consequence of people viewing it (as well as its continued availability to be viewed). Harm suffered may take the form of a slight to the person's reputation, even in the case of those who are deceased (e.g., Whitney Houston, Marilyn Monroe, but also noncelebrities) (see Aristotle, 1999; Feinberg, 1984; Pitcher, 1984, for further discussion on harming those who are deceased), or psychological harm (i.e., experiencing shame, even when unwarranted, or embarrassment, or lower self-esteem, anxiety, and depression). The deepfake$_p$'s notoriety may cause the subject to suffer professionally in terms of future career prospects, or at least increase the risk of this happening. It does, however, remain an unanswered empirical question exactly how much impact the deepfake$_p$ would have on one's reputation, career, and psychological health, or how much public knowledge of the fact that the sexual activity depicted is not genuine would mitigate this impact. Despite the paucity of research on the topic, anecdotal evidence is available to shed some light on the matter.

Henry et al. (2018) report the case of an Australian woman whose photographs were stolen from her social media accounts, used to create deepfake pornography, and then posted on numerous websites. The experience, we are told, caused her to feel "physically sick, disgusted, angry, degraded, dehumanised" (p. 1). Celebrity Scarlett Johansson, whose face can be found on a number of deepfake$_p$ videos, had this to say about her situation: "Clearly this doesn't affect me as much because people assume it's not actually me in a porno, however *demeaning* it is" (Harwell, 2019, p. 1; emphasis added).

The Australian woman is reported to have experienced harmful effects; Scarlett Johansson, in turn, describes the deepfakes$_p$ featuring her as demeaning, even though she is of the opinion that people will tend to assume that the videos featuring her are fake. As a further example, NM, who is not a celebrity and who we are told has faced over six years of deepfake$_p$ harassment—for initially speaking out against deepfakes$_p$—"doesn't have the benefit of people knowing that these images must be fake[,] as they would assume with famous actors" (Curtis, nd, p. 1). Moreover, the fact that "these deepfakes are still easily found by searching [her] name, raises questions about her future employability and online reputation" (*ibid.*).

The fact that deepfakes$_p$ are mostly targeted at women can also be viewed as further evidence of "a cultural climate in which women are systematically treated as unequal, and this inequality is achieved in large part by treating women as sexually unequal" (Patridge, 2013, p. 29), and therefore as further evidence of the *harm* women have to endure as a consequence of this inequality.

DEEPFAKES$_P$ AS INDICATIVE OF A LACK OF RESPECT

Finally, irrespective of any harm caused, or even the increased risk of harm, deepfakes$_p$ are morally problematic because of:

4. The disrespect they show to the target of manipulation

Deepfakes$_p$ are indicative of a failure to treat the target of manipulation as someone deserving of respect.[6] In Kantian terms, the intent behind the creation of the deepfake$_p$ violates the second formulation of the categorical imperative whereby one should "act in such a way as to treat humanity, whether in your own person or in the person of another, always at the same time as an end and never simply as a means" (Kant, [1785] 1993, §429).

To be clear, it is not my contention that all deepfake manipulation is necessarily disrespectful and therefore necessarily violates Kant's categorical imperative. Take the Buscemi-Lawrence deepfake as an example. Actor Steve Buscemi's face is manipulated so as to replace the face of actress Jennifer Lawrence when speaking at the *Golden Globe Award Ceremony*.[7] The first thing to note about this deepfake is that it is unlikely that its creator sought to pass the deepfake off as genuine footage of Steve Buscemi wearing a dress and impersonating Jennifer Lawrence at the *Golden Globes*. There seems to be no intent to misrepresent without disclosure (although, admittedly, the disclosure is not explicit). Second, it seems reasonable to interpret the deepfake as a piece of nonmalicious entertainment not intended to be disrespectful, and/or as an aesthetically interesting, certainly novel, creation.

As reasonable as this interpretation is, however, I am willing to concede that some (perhaps many, including Buscemi and Lawrence) would disapprove of the deepfake, even though it is nonpornographic (see, for example, MacDonald & Palmer, 2019). It is not inconceivable that Buscemi and Lawrence would object to being the butt (as they perceive it) of this particular deepfake joke owing to the (alleged) lack of respect it shows.[8] That said, it is worth noting that Steve Buscemi seemed more bemused than offended, or otherwise distressed, by the clip when he was first shown it on a TV talk show (I am unaware of Jennifer Lawrence's reaction, assuming she has seen the clip). But even if this or other objections are valid, it is still conceivable that *some* deepfake manipulations are sufficiently benign as to not be construed (in any reasonable sense) as disrespectful. While this may be true, it is my contention that pornographic deepfakes are *not* among them—a claim I will now defend.

To disrespect someone—irrespective of whether they are a celebrity like Scartlet Johansson or lesser known like NM, and irrespective of their gender—is to dismiss them or some aspect of them—for example, their thoughts or feelings—as irrelevant; it is to impute on them a lesser value (Hill, 1987). Showing respect, however, requires not only that one refrain from this type of behavior but also that one treat all individuals as persons and, on account of their standing as persons, tailor one's actions and attitudes (thoughts/feeling) to accord with this standing. Disrespect, Eidelson (2015) notes, "is the absence of appropriate responsiveness to someone's standing as a *person*" (p. 7; emphasis added).[9] For Eidelson, then, "what is of fundamental importance to determining whether an action is disrespectful . . . is the set of reasons for which the agent acts, rather than the effects as social meaning of her action" (*ibid*.).[10] In other words, for an act to be disrespectful, *irrespective of any harm caused*, the perpetrator of the disrespect must intentionally fail to do what we are (and should be) obliged to do: namely, act in accordance with the intrinsic value the person possesses in virtue of being a person.

Disrespect is not therefore injurious simply because it places constraint on an individual's ability to act or because it necessarily causes physical, emotional, or some other form of psychological harm, although an act of disrespect may do any and all of these things (see, for example, Sokol-Heener et al., 2015). In addition, disrespect is injurious (to a greater or lesser degree) whenever it impairs the target of disrespect in their pursuit of a positive understanding of themselves (Honneth, 1992): for where respect acts to reassure the individual that they are recognized as a person and valued because of it, disrespect seeks to take this away. But even if one's act of disrespect should fail in its intent to be injurious in the ways described, it remains morally problematic because of its *intent*. Moreover, as is the case with harm, given the propensity of deepfakes featuring women, the intent to be disrespectful through the use of this medium is further evidence of a more universal lack of respect for women.

Irrespective of one's particular view of nonpornographic deepfakes like Buscemi-Lawrence, and irrespective of the *conceivability* of nondisrespectful deepfakes,[11] deepfakes that constitute nonconsensual pornography are disrespectful if they are created as a means of expressing one's disregard for another person as a person of value, and therefore as a means of dismissing, through one's action, that which we are morally obliged to do: namely, recognize and respond appropriately to the inherent value we each have as persons. To illustrate: it is a contingent fact that, for many people, sexual activity, and certainly their own sexual activity, is an intensely private matter. The intentional, nonconsensual violation of this privacy, including the *appearance* of such a violation, as deepfakes$_p$ are prone to provide, irrespective of disclosure (I would argue), amounts to a means of disregarding a contingent fact about our values. Recall that such disregard is injurious if it impairs our pursuit of a positive understanding of ourselves (Honneth, 1992). It is difficult to reconcile the *demeaning* nature of nonconsensual pornographic deepfakes (as Scarlett Johansson describes them), or the feeling of being *degraded* or *dehumanized*, as was also reported, or the sense of shame, humiliation, or embarrassment one might experience on discovering one is the "star" of such a video, with the pursuit of a positive understanding of oneself. Where one *intends*, through one's actions, to impact negatively on another person in this way, then one has acted disrespectfully, irrespective of whether one succeeds in achieving this aim.

CONCLUSION

Pornographic deepfakes (deepfakes$_p$) are morally problematic for at least the reasons discussed: the lack of consent inherent in nonconsensual pornography, their potential to cause harm, and their use as an expression of disrespect. Other types of deepfakes may be vulnerable to a similar charge; whether they are or not needs to be considered on a case-by-case basis, however. That issue aside, I have restricted my interest to deepfakes$_p$ because, as the most prevalent form of the deepfake video, they constitute the most popular use (or perhaps abuse) of this particular technology and therefore warrant closer moral scrutiny.

I will return to the issue of deepfakes in the next chapter when contrasting them with the chapter's focus: private sexual fantasy.

NOTES

1. A further moral and legal concern may be how the images used to create the faceset were accessed: Were they all in the public domain, for example? This may be particularly pertinent in the case of deepfakes used as revenge porn. Without wishing to diminish the importance of this issue, I will not pursue it here.

2. https://deeptracelabs.com/mapping-the-deepfake-landscape/ (accessed October 11, 2019).

3. In addition to FakeApp, the *DeepNude App* is designed to create a nude image of a person (again, usually a woman) from an original, partially or fully clothed, digital photograph or similar image (Vincent, 2019).

4. For this reason, Delfino describes the original porn actor as a victim of the deepfake$_p$ deception, also.

5. They also misrepresent the sexual activity of those in the original recording, but this is not the focus of our discussion, here.

6. The same lack of respect is arguably directed at those who appear in the original pornographic footage although, again, I will focus only on those whose faces have been inserted rather than replaced.

7. https://www.youtube.com/watch?v=r1jng79a5xc (accessed October 15, 2019).

8. I am not suggesting that any act of ridicule violates Kant's categorical imperative. One may have acted in a way deserving of ridicule, for example. Ridicule as warranted criticism is to treat the ridiculed individual as an end in themselves and not merely as the object of one's ridicule.

9. I emphasized the word "person" because I wish to distinguish the immorality of disrespecting S in virtue of not recognizing or dismissing their value *as a person* (a human being) and disrespecting, inter alia, the social status, job, affiliations, opinions of that person. Disrespect in the latter regard, unless judged to be justified (i.e., disrespecting, the social standing of a tyrant, or the job of a contract killer), *may* be regarding as a lesser (suberogatory) wrong, including an act of poor taste.

10. See also Darwall (1977) and Waldron (2012) for detailed discussion on respect and dignity.

11. It is conceivable that all parties involved in the pornographic deepfake (e.g., the original actors and the target of manipulation) consented to the manufacture and distribution of the deepfake$_p$ as pornography and, as such, do not consider the creator's intent to be disrespectful; nor (let us allow) was his or her intent to be disrespectful. Were this to be the case, then the deepfake$_p$ would not be an example of *nonconsensual* pornography.

Chapter 12

Fantasy and Fictional Immorality

INTRODUCTION

In this chapter I examine a means of depicting fictional immorality that I have thus far avoided: namely, private fantasy. The reason for this deferment, you may recall, is that private fantasy is not available to others unless presented in a form that subscribes to my original liberal definition of fictional depiction. Thus, when S fantasizes about having an extramarital affair with the cartoon character Marge Simpson, while this private fantasy *depicts* fictional immorality (*qua* infidelity) in the form of a mental image (let us allow), it is not a fiction that is depicted using a more publicly accessible medium: a drawing or painting, story, diary entry, or song, or any other physical construction, including behavior.

The term "private fantasy" is not therefore tautological, for fantasies are not *necessarily* private. Instead, and as we have seen when discussing fiction more generally, through various modes of depiction (e.g., literature or film, or whichever medium one selects), one can share fantasies with others and even create a collective fantasy among a group of likeminded enthusiasts. Private fantasies, in contrast, describe fantasies that are solitary pursuits involving an author and audience who are one and the same (Cooke, 2014). When referring to the fantasy as "private," then, I do not mean simply that the depiction is kept away from others (under lock and key, as it were) or that it is in some other way confidential; rather, I mean that it is not directly accessible by another. Contrast this type of depiction with the occasion on which S *enacts* the content of their private fantasy while playing the central (fictional) character in a self-penned play about sexual fantasy. When presented in this way, the enacted content adheres to the requirements of the more liberal interpretation of depiction I have until now employed throughout the monograph.

There is another reason for my eagerness to treat private fantasy as a discrete category of fiction. It may be that S wishes to fantasize about his relationship with Marge Simpson in a not-too-dissimilar way to how one might imagine Garth, from the 1992 film *Wayne's World*, fantasized about Bugs Bunny prior to asking his friend, Wayne: "Did you ever find Bugs Bunny attractive when he'd put on a dress and play a bunny girl?" Nevertheless, it is perhaps more common for fantasies, particularly of a sexual nature, to involve actual people (Joyal et al., 2015): whether friends or family, mere acquaintances or strangers we pass in the street or sit across from us on public transport or in cafés, or famous personalities we will likely never meet. In each case, while the fantasy is a fiction insofar as it represents a counterfactual or hypothetical event (hereafter, counterfactual)—say, some consensual or nonconsensual sexual liaison that has not happened—it is not one that involves fictional characters and is therefore not in keeping with my definition of pure fictional immorality.

The fact that private fantasies are likely to feature depictions of *actual* people in some form of counterfactual event is a point that should not be ignored or dismissed, of course. Prior to discussing the moral implications of this likelihood, however, I wish to continue with the approach to fiction I have been using throughout the majority of this book and discuss private fantasy in the context of pure fictional immorality, before moving on to consider the moral implications of fantasizing about actual persons (see section "Fantasy and Matters of Consent, Harm, and Respect" and onwards).

In this chapter, I assess the morality of *private* fantasy as a distinct category (*qua* depictions that remain confined to one's head, so to speak). In particular, I am interested in whether the fact that one's private fantasies are only for one's own consumption protects them from certain, perhaps *all*, moral admonishment regardless of content, meaning, and motivation. In short, should the privacy element, unique to this particular means of depicting fictional immorality, make moral concerns about private fantasy (hereafter, simply "fantasy") a nonstarter? I do not believe it should, for reasons I will discuss.

Before engaging with this matter further, however, I present a number of moral concerns people have had over the years with the act of fantasizing, beginning with Cherry's (1988) argument in favor of the immorality of fantasy in virtue of its self-gratifying and usurping nature.

USURPING AND SELF-GRATIFYING FANTASIES

For Cherry (1988), the immorality of fantasy—its invariable wrongness, as he describes it—stems from the kind of activity fantasizing is. As he states: "There must be something about the way in which, and purpose with which, ideas are brought and held before the mind which makes the very activity of fantasising a corrupt exercise of consciousness" (p. 126). The seed of fantasy's immorality lies, then, in the fact that it (the *act* of fantasizing) "usurps finer possibility" (*ibid.*, p. 129). In fact, even the most mundane and innocuous activities can be replaced by fantasy (Hershfield, 2009), not just those with more moral weight. In short, should I do good or merely fantasize about it? Should I *intend* to do good or merely fantasize about it? On either count, fantasizing is standing in for a better act (including the act of intending to act). That said, Cherry does acknowledge that doing or intending may not always be possible. After all, in extreme cases, perhaps fantasy is all that one has available.

Consider Terry Waite, special envoy of the Archbishop of Canterbury in England, who, while trying to secure the release of four captives, was himself held hostage in Beirut for nearly five years (1,763 days, between 1987 and 1991). For the first four years he was held in solitary confinement and, for part of that time, chained to a radiator twenty-three hours a day.[1] Imagine someone in the same situation as Terry Waite. Given that they are extremely limited in the good they can do or even intend to do—unless one is referring to some long-term intent that perhaps stems more from hope than anything else—should this person at any time during their captivity fantasize, is it appropriate to decry their activity as immoral? If fantasizing is immoral because it usurps doing good or intending to do good, then where it is not possible to do either of these things, the act of fantasizing is not immoral, at least for *that* reason.

Leaving such extreme cases aside, is Cherry correct to present "usurping a better activity" as sufficient grounds for deeming fantasy immoral? Perhaps if one spent all or a good part of one's time fantasizing to the detriment of any actual good—akin to the protagonist, Billy Fisher, in Keith Waterhouse's novel *Bill Liar* or Walter Mitty in James Thurber's *The Secret Life of Walter Mitty*—then Cherry may have a point. But to make the activity itself immoral, and therefore something one ought not to do at all, places unrealistic expectations on how we ought to live our lives and leaves it unclear why the same prohibition should not be placed on all nonessential activities, at least where they have the power to usurp positive undertakings.

More importantly for Cherry, however, the inherent wrongness of fantasy is grounded not on the act of usurping finer possibility but on the fact (as he sees it) that "[w]hat prompts a person to fantasise is, broadly speaking, self-gratification; and this precisely cannot be a motive to disinterested benevolence" (1988, p. 132). According to Cherry, then, the content of the fantasy, no matter how mundane or innocuous, is characteristically self-centered, making the act of fantasizing self-absorbed and self-congratulatory.

Hershfield (2009) questions Cherry's characterization of fantasy. Is it really the case that all fantasy is as self-referential as Cherry claims? Even if it is, why does this make fantasy (the *act* of fantasizing) *immoral*? Either the problem is the usurping nature of fantasy, which presumably would remain the case irrespective of any self-referential content (if we deny that the content of fantasy is *necessarily* self-referential), or it is the self-referential content itself, which seems independent of the *amount* of time one spends fantasizing, or the usurping credentials of fantasy. Yet, in the latter case, it is not clear why one should be morally obliged not to indulge in a little self-referential fantasy.

I find Cherry's argument unsatisfactory, for the reasons discussed, and concur with Hershfield: that Cherry fails to provide sufficient detail to support his universal condemnation of fantasy. Let us move on.

ADDITIONAL MORAL CONCERNS

Additional moral concerns over fantasies (that depict fictional immorality) take the following form:

1. Minimally, engaging in fantasies (of this kind) increases the risk of actualizing them, thereby causing harm.

We have encountered a similar concern to this already when discussing fictional immorality more generally (e.g., media violence in which one reads, watches, or listens to a fiction featuring a character who has nonconsensual sex with a stranger, or where one actively engages in this act through one's avatar in a video game). On these occasions, concern was grounded on the harm principle, and evidence of harm was sought in the form of a posteriori discoverable associations between one's engagement with the fiction and the likelihood that it will cause harm or increase the risk of harm. The same approach is being applied, here, to fantasy: that fantasizing about rape makes it more likely that one will (attempt to) rape someone (for discussion on research findings in relation to this issue, see section "The Morality of Idle Fantasy"). Furthermore:

2. Where the fantasy is an expression of one's beliefs and/or desires, in addition to increasing the risk of actualizing the fantasy and therefore causing harm (in accordance with #1), the beliefs and/or desires themselves are immoral.

Fantasizing about having nonconsensual sex with a fictitious stranger not only requires greater agency in generating the fiction than reading a book featuring rape or viewing a film or listening to a song with similar content, more than this—perhaps even because of it (the argument goes)—the fantasy is an expression of a set of preexisting beliefs and/or desires (e.g., the belief that women enjoy forceful sex and that "no" really means "yes" or simply that one desires to engage in nonconsensual sex, irrespective of one's beliefs about it).

Again, a similar move can be found in chapter 5 when discussing the supplementary harm conditions. In chapter 7, a distinction is made between enjoying the simulation, in and of itself, and using the simulation as a vehicle to elicit enjoyment for what one really enjoys: that which the simulation is a depiction of. In the sections "The Morality of Idle Fantasy" and "The Morality of Surrogate Fantasy," I apply a similar distinction to fantasy and examine the moral implications of such a move.

3. Irrespective of the increased risk of actualizing the fantasy (and therefore increasing the risk of harm), and irrespective of whether one holds beliefs and/or desires congruent with that fantasy (in a manner yet to be discussed), if the immorality depicted is directed toward or involves other *actual* people, it is immoral because it disrespects the individuals depicted therein.

According to this "disrespect" objection, fantasizing about that which is immoral, at least where the immorality is directed toward or in some way involves another person, is immoral irrespective of any *direct* harmful consequences (or the increased likelihood of these) and irrespective of whether the fantasy stems from beliefs and/or desires (of a kind to be discussed), because, a priori, such a fantasy demonstrates a lack of respect to those who feature in the fantasy. Because of this, it amounts to an indirect form of harm. The "disrespect" argument cannot, of course, be applied to fictional characters. Therefore, where a fantasy features fictional characters/events exclusively (perhaps with the exception of oneself), the applicability of the "disrespect" objection to actual persons (members of wider society) who do not feature in the fantasy is in need of greater defense (i.e., a greater defense of the idea that fantasizing about Marge Simpson or Jessica Rabbit is disrespectful to actual women is required). I will have more to say about this objection in the sections "Fantasy and Matters of Consent, Harm, and Respect" and "Fantasy," both in terms of fantasy featuring actual persons and those that do not.

FANTASY'S AUTONOMOUS NATURE

The Absence of a Direction of Fit

Returning to point 2, Hershfield (2009) asks how our fantasies are functionally related to our desires and cognitive states—such as beliefs, attitudes, and intentions—and ultimately our behavior. Borrowing from Searle (1983), he argues that for a belief to be satisfied, it must be a "good fit" with some corresponding state of the world, whereas fantasies are not dependent on the world in this way (in terms of the way the world is or ought to be). Thus, for my belief that there is a table in front of me to be true (for this belief-state to be satisfied), there must be a table in front of me. A true belief therefore has mind-to-world direction of fit, insofar as it represents the world as being a certain way that corresponds to the state of the world at that time. A desire, on the other hand, has world-to-mind direction of fit. The desire represents how the world needs to change in order for the desire to be satisfied. If I desire a bigger apartment than the one I currently occupy, then the world needs to change, in a manner that corresponds to the content of my desire for the desire to be satisfied.

As a fantasy does not have a direction of fit with the world, the world does not have to be a certain way or need to change to something else for the truth of the fantasy to be satisfied (for it to be true that what I am depicting is a counterfactual event). The act of fantasizing is sufficient to satisfy the truth of the fantasy. I do not have to live in a bigger apartment for it to be true that my fantasy depicts me living in a bigger apartment (and for living in a bigger apartment to be a counterfactual event, given my current accommodation). Neither is it necessary for the world to be different (to be a place in which I would have a bigger apartment) in order to indulge my fantasy. Instead, my fantasy can be satisfied (*qua* a fantasy) irrespective of the state of the world and therefore the directions of fit required for the satisfaction of my beliefs and desires.[2]

The content of fantasy is not intended, necessarily, to represent how things are or how they ought to be. Our fantasies therefore have the capacity to be independent of our beliefs and desires. Consequently, "the implicit rules of fictive imagining preclude that we infer anything about the real-world" from fantasy (Bartel & Cremaldi, 2018, p. 39). This means that nothing about the real world need be inferred from the fact that (in a manner reminiscent of Seth Grahame-Smith's *Abraham Lincoln: Vampire Hunter*) I fantasize that Winston Churchill was a demon hunter when he was not governing the United Kingdom or that Queen Victoria, in her later years, moonlighted as a Punch and Judy puppeteer at the end of Clacton Pier. As a consequence, I can be well acquainted with the lives of both Churchill and Queen Victoria without this knowledge interfering with my ability to indulge either fantasy.

Involuntary and Voluntary Fantasies

Declaring that one's fantasies do not have to align with one's beliefs or desires is not to declare that they never do and therefore that they are never an expression of one's beliefs and/or desires; rather, it is to acknowledge that they (fantasies) do not have to be. Independent of this relationship, fantasies may also be voluntary or involuntary.

Involuntary fantasies (or what Smuts, 2016, calls *spontaneous* fantasies) can be likened to unbidden thoughts insofar as the emergence into consciousness of their content is outside of voluntary control. The image of an alluring Marge Simpson or Jessica Rabbit may suddenly "pop" into my consciousness. It may even defy my will not to experience it, as annoying tunes heard on the radio or in a supermarket (and so on) are prone to do every now and then: for, as Neu (2002) notes, "our control over our thoughts may be far more limited than our control over our actions" (p. 141). Given this, and where ought implies can, on such an occasion, I ought not to be held morally responsible for the occurrence of this type of *passive* fantasy and, instead, be regarded as blameless (Bartel & Cremaldi, 2018).[3] James (1993), however, offers a caveat:

> Although many fantasy images seem to enter consciousness unbidden, the decision to dwell upon them is usually voluntary. Such decisions to fantasise, insofar as they are voluntary and intentional, are as subject to ethical assessment as any other human choices. (p. 51)

James' point is that even if an unbidden thought did pop (involuntarily) into my mind, the act of *entertaining*, perhaps even *embellishing*, it, as opposed to, say, trying to dismiss it, *is* under my control and, as such, constitutes a legitimate target for moral scrutiny. Bartel and Cremaldi (2018) express a similar view when distinguishing between passive fantasies that arrive unbidden, which may not reflect our desires, and active fantasies, which take the form of "imagined scenarios that we return to time and time again . . . [which we] deliberately tell ourselves . . . , develop, refine, and relish" (p. 42). Moreover, for Bartel and Cremaldi, not only do these active fantasies track our desires, they help cultivate them. Consequently, "if it is morally wrong to desire *x* then surely it would be morally wrong to cultivate a desire for *x*" (*ibid.*, p. 44). It is, of course, also important to note that the moral implications of distinguishing between voluntary and involuntary fantasy become something of a moot point *if* it can be shown that *all* fantasies are morally innocuous (Hershfield, 2009). Whether they are or not, I contend, turns not just on their voluntary status but on whether they are *idle* or *surrogate* fantasies.

Idle and Surrogate Fantasies

Leaving passive fantasies aside, owing to their lack of (conscious) control, are all *active* fantasies morally innocuous? When I say "all," it is, of course, important to reiterate my interest in depictions of fictional immorality. *Pace* Cherry (1988), I do not share the view that fantasy per se is immoral (for the reasons discussed); therefore, my question concerns fantasies that depict fictional immorality, only. Are *these* active fantasies morally innocuous? To answer this question, we must first distinguish between active fantasies (again, hereafter, just "fantasies") that are *idle* and those that are *surrogate* (Cherry, 1988). Before proceeding, however, a point of clarification. In keeping with what has been discussed thus far, reference to idle or surrogate fantasies should be taken to mean *private and voluntary* idle or surrogate fantasies.

In addition to presenting as counterfactuals conditions, fantasies will more likely than not feature the agent (even if it is not in an exhaustively self-gratifying manner). He or she will fictively imagine themselves engaging in an action, or as somehow involved in an event, that has not happened, especially in cases in which the fantasy involves fictional characters (because what is depicted cannot actually happen to them, of course). By engaging with the fantasy—*qua* entertaining and embellishing the fictively imagined counterfactual action/event—does the agent desire the actualization of the fantasy's content and therefore the counterfactual event? In other words, given that the fantasy is private and voluntary, does it track the agent's desire (namely, that which the fantasy depicts) or is it, instead, that the object of the agent's desire is simply the fantasy itself (the depiction *qua* depiction)?

Regardless of how convincing the representative quality of the fantasy is, should we, as Cherry (1988) asks, assume that we fantasize *only* about those things we desire because we desire them, or does fantasizing serve some other purpose? If the former, then using the fantasy to satisfy vicariously that which is desired (the *desideratum*) means, necessarily, that what is depicted within the fantasy is a *depiction* of the desideratum, not the desideratum itself. If, however, the fantasy serves some other purpose, then this has yet to be determined. One suggestion, that I intend to pursue, is that the desideratum is found only within the fantasy, *qua* some internal property, and is not therefore external to it, as would be the case if what was desired were only referenced by the fantasy rather than being the fantasy.

For Cherry, a fantasy that acts as a "[substitute] for an external—pre-existing and presupposed—desideratum" (1988, p. 118) or as a *faux de mieux*[4] for the real thing is surrogate in nature. One that is active rather than passive and is *not* tracking a desire for anything beyond the fantasy itself—and therefore seeks only to indulge the fantasy itself and nothing more—is, in contrast, an idle fantasy. Thus, "there is no prior desire to do, no inclination to actualise, states of affairs answering to [idle] fantasies" (Cherry, 1988, p. 119).

THE MORALITY OF IDLE FANTASY

Recall that "fantasies" (as I am using the term here) are solitary pursuits involving an agent who is also the audience. As a consequence, it is tempting to think of the agent, within this dual role, as creating for themself a representation of what they desire (i.e., if I fantasize about murdering random strangers, then I must desire to murder or in some sense approve of the idea of murder, thereby making my fantasy a direct window into my desire). While it may be tempting to make this connection, and while it may even have a certain intuitive appeal, there is no entailment (logically or conceptually) between one's fantasy and one's desire or belief/pro attitude (i.e., one's fantasy is not necessarily an expression of one's desire and/or approval of what is depicted). As Kershnar (2005) notes: even where fantasizing about rape is arousing for the individual concerned, this does necessitate that he or she wishes to rape or be raped. Nor does empirical evidence support the contingent truth of a connection, on all occasions, between one's fantasy and one's belief/pro attitude and desire (recall points 1 and 2 from the section "Additional Moral Concerns").

Hershfield (2009), for example, discusses a study by Laumann et al. (1994) in which those surveyed reported that they were disinclined to act on their fantasies about having an affair, suggesting that fantasy is not necessarily action determining. In addition, Cooke (2014) notes that rape fantasies are common among men and women, yet there is no compelling evidence of a connection between rape fantasies in men and a disposition to rape or hold positive beliefs or attitudes about rape, or hold that women want to be raped or be of the opinion that women on the whole desire this (see also Critelli & Biovana, 2008; Biovana & Critelli, 2009). An idle fantasy about rape or murder, or whichever immoral act it may be, evidence suggests, is capable of being an autonomous fantasy that stands apart from, and is therefore independent of, the agent's desire (although see Bartel & Cremaldi, 2018, below for a different view).

To further support the claim for autonomy, although I concede it involves a degree of extrapolation, consider the subcategory of criminal activity referred to as *sexual voyeurism*, which describes an individual who views child pornography (sometimes compulsively) in the absence of a documented history of sexually soliciting minors (and with apparently no interest in doing so). When the computers of *sexual voyeurism* offenders are confiscated, there is no evidence of sexually inappropriate "chats" with children and, following publicity after their arrest, no reports of children, including their own, coming forward accusing them of any form of inappropriate contact (Berlin & Sawyer, 2012). Given this, Berlin and Sawyer conclude:

> Some individuals appear to be experiencing compulsive urges to voyeuristically view such images [of child pornography], devoid of any motivation to actually approach a child sexually. In other words, in such instances, the act of voyeuristically, and often compulsively, viewing such imagery over the Internet would appear to be an end in and of itself; rather than a means to some other end—such as actual sexual contact. (2012, pp. 31–32)

Sexual voyeurism is a criminal offense, and given the nature of the offense, it should be morally condemned. Importantly, though, while it is more often the case that those with a sexual interest in children *do* possess, or are motivated to possess or at least view, images of children that violate child pornography law (Kingston et al., 2008; Riegal, 2004; Seto & Eke, 2005), what the category of sexual voyeurism indicates is that viewing these images can nevertheless occur independently of any desire (it would appear) to act inappropriately toward a child (beyond viewing inappropriate images of them, that is).

For those classified as having a compulsive desire to repeatedly view sexually explicit images of (prepubescent)[5] children, the primary components of their paraphilic disorder, Berlin and Sawyer (2012) inform us, are voyeurism and pedophilia. The pedophilic aspect is restricted to voyeurism over the Internet and, as such, should be distinguished from pedophilia per se, whereby an individual is motivated to *engage* in a sexual act with a prepubescent child (characteristic of *direct victimization)* and not simply view images of young children with sexualized content. That said, recall from chapter 3 that the images of children used to elicit sexual arousal may not depict sexual abuse; they may not even contain images of children completely or even partially nude. In and of themselves, the images may be quite innocuous and not violate child pornography law (Tate, 1990). As Howitt (1995) points out, sexual stimulation may not be based on explicit sexual content but on the *fantasy* occurring in the mind of the offender. Given this, if one were to remove the criminal element from *sexual voyeurism* (in the form of the lascivious viewing of images of child sexual abuse *qua* child pornography)—in fact, if one were to remove any physical depiction at all—and if, in the absence of this direct physical stimulus, the sexual voyeur merely fantasized about children in a sexual way, without his (or her) private activity escalating into sexual solicitation or molestation, and was content to do this, then it would seem that we have a further example of idle fantasy, and the separation and autonomy of fantasy from desire. Of course, one might object to such a fantasy. If the objection is to stand, however, in the absence of broadening its scope to include rape fantasies and possibly *all* fantasies featuring depictions of immoral content, and therefore in the absence of support for the universal condemnation of fantasies featuring depictions of immorality, it needs to articulate what is peculiar about *this* (pedophilic) type of fantasy to warrant special moral attention.

In response, one could question the likelihood of this type of fantasy being a genuine idle fantasy, and therefore treat with skepticism the idea that the content really is divorced from the desire. To be clear, such a move would not have to deny the ontological independence of idle fantasy (from desire and/or attitude); rather, it would only have to question how likely it is that the level of separation characteristic of idle fantasy was actually operating in an individual who fantasized about child sexual abuse or other depictions of immorality, such as rape. As Bartel and Cremaldi (2018) caution, while it remains true that one can indulge one's fantasy without endorsing its content, and, from this, true that an individual may desire to fantasize about something they do not desire for real (in accordance with Cooke, 2014; Kershnar, 2005; see also Grimshaw, 1993; Hunt, 1974; Neu, 2002), we nevertheless "reject the stronger claim that acts of fictive imagining are *always* sufficiently bracketed out from the agent's real-world beliefs or desires" (p. 45; emphasis added). Because of this:

> When an agent engages in an act of imagining (whether fictive or otherwise) in order to *cultivate* an immoral desire, then that agent's act of imagining should be open to moral scrutiny as their imaginative engagement is not genuinely bracketed out from their real-world desires. (*ibid.*; emphasis added)

Fantasizing as a means of cultivating an immoral desire is a feature of surrogate fantasy. Here, the content of the fantasy is congruent with one's beliefs and/or desires because it is a means of expressing them and, in the case of desire, satisfying it, vicariously.

THE MORALITY OF SURROGATE FANTASY

Idle fantasies are characterized by the autonomous nature of the fantasy, which is capable of operating independently of one's desires, core beliefs, and values; and as we have seen, the difference between a fantasy about having an extramarital affair and a desire to engage in one is that the latter has content that represents a changed world: one in which the world ought to be different so as to accommodate—to satisfy—the desire. In the case of fantasy, no such change is sought or required to accommodate the pretense. As we have seen, the autonomy of fantasy is a contingent fact about fantasy not a necessary one, meaning that some fantasies may express the beliefs and/or desires of the agent. The fact that one fantasizes about having an extramarital affair could therefore be because that is what one desires and, in the absence of satisfying this desire, at least directly, or even as a precursor to its actualization, one can obtain vicarious satisfaction through a surrogate fantasy. As Neu (2002) notes:

> Fantasies in themselves may do no harm, even if their content is appalling rather than ultimately innocent. I have said it is the underlying desires, not the fantasies, that lead to action. But it must be acknowledged and emphasized

> that fantasies like symptoms may reflect potent desires rather than ineffectual wishes. (p. 154)

In keeping with this view, for Cooke (2014), fantasies (in the form of fictive imaginings) should not be considered morally problematic because, in and of themselves, they do not require nor necessitate alethic commitment. As we have seen, fantasies lack a direction of fit. Morally troubling fantasies are possible, however, when they express the agent's underlying desire. What may not be uniform and therefore clear in individual cases is the direction of cause. Thus, while it may not always be the case that the fantasy originates as an expression or symptom of the agent's underlying desire (one direction of cause), as Kershnar (2005) points out, sexual fantasy, even if at first idle, may *lead* to nonautonomous changes in the agent and therefore what the agent desires (a different direction of cause). Whether these changes occur or not remains a yet-to-be-resolved empirical matter. Nevertheless, Bartel and Cremaldi (2018) argue, in the case of idle fantasy, that there is a risk that one may be tempted to develop (in an intentional sense), or come to acquire (unintentionally, as Kershnar warns), beliefs and desires congruent with the content of one's initially autonomous fantasy.

For this to occur (intentionally or otherwise), as we have seen, beliefs and desires must have object-to-world or world-to-object direction of fit. One must therefore come to believe, in a manner conducive to one's fantasy, that women enjoy being raped (for example) or, in a manner conducive to one's desire, that the world ought to be a place in which I am able to contribute to their enjoyment of rape. Of course, it may not always be easy to "[distinguish] between fantasies which serve as harmless safety valves and fantasies which should be taken as symptoms, as warning signs" (Neu, 2002, p. 155). Nevertheless, it is important distinction to make, conceptually, as this distinction informs (and should inform) our moral judgments.

In the case of idle fantasy, even if we concede that it is illegitimate to conclude, from the simple act of fantasizing about rape, that the agent engaged in the fantasy desires to rape or be raped, or approves of rape, it nevertheless remains reasonable to infer, at least in the case of an erotic, as opposed to an aversive or a sadistic, rape fantasy (see Biovana & Critelli, 2009[6]) that the same agent must be fictively imagining that rape is enjoyable for the person being raped *in the fantasy* (see also Bartel & Cremaldi, 2018). In the case of someone—a man—fantasizing about raping someone else—a woman, let us say—for the act of being raped to be (fictively imagined as) enjoyable, the woman depicted in the fantasy, even in the case of a fictional character (e.g., Marge Simpson or Jessica Rabbit), must bear some resemblance, psychologically, to an actual woman. Given this, there is a risk that the idea of enjoyment will become associated with any act of rape, not just the one fantasized, and play a causal role in the transference of the unrealistic outcomes of fantasy to real life. Borrowing from Gaut (1998), it is arguably the case that for the one-time idle fantasy to promote actual beliefs and desires about rape (i.e., become a surrogate fantasy), rather than their mere pretense and/or a desire for the pretense, given that the newly acquired belief and/or desire was not a feature of the fantasy in its original form, the women who are depicted as enjoying rape in the original idle fantasy must resemble, in some significance way, real women. Kershnar (2005), however, suggests that, as fantasies are specific, one must be cautious about generalizing to all cases of rape and therefore guarded about how much agents infer a general claim (about women, for example) from their specific example within the fantasy.

In the end, then, any potential harm caused by fantasy, either directly in the form of behavioral change or indirectly through a change in one's beliefs/attitudes/feelings, awaits empirical support. As is the case with depictions of fictional immorality more generally (see chapter 6), there is a lack of consensus on what the effects of engaging with this type of fiction are.

FANTASY AND MATTERS OF CONSENT, HARM, AND RESPECT

Recall the discussion on pornographic deepfakes (deepfakes$_p$) found in the previous chapter. As a precursor to discussion on fantasy and matters of consent, harm, and respect, consider the difference, morally speaking, between one's fantasy about having nonconsensual sex with a celebrity and creating a fictional depiction of the same encounter as a deepfake$_p$ that is stored away for personal use only (Öhman, 2020, draws a similar comparison). There is a clear physical difference between the two, of course; but is this sufficient to establish a moral difference? Moreover, in both cases, the fictional content features depictions of *actual* people. In the deepfake$_p$, even if I exclude my own manipulated face and the bodies of the original actors, at the very least, there remains for moral consideration the manipulated face of the celebrity; and as the celebrity also appears in my fantasy, in neither case are we discussing examples of pure fictional-x_i.

Private Deepfake

When comparing the two examples, let us allow that the private deepfake$_p$ (hereafter, deepfake$_{private}$) was created without the consent of either the celebrity or any of those featured in the original footage. Let us also contrive that I was the only one involved in manufacturing the deepfake$_{private}$; only I have access to it, and I view it only once before destroying it. I therefore feel secure in the knowledge that no one else has seen the deepfake$_{\text{private}}$. Finally, I never mention what I have done to anyone.

Given the fact that neither the celebrity nor the original cast members consented to the creation and use of the deepfake$_{\text{private}}$, the lack of consent remains an issue. My deepfake$_{\text{private}}$ is still a species of nonconsensual pornography, subject to the same moral concerns discussed in the previous chapter.

Focusing, now, on the question of harm. What is less clear, *in this case*, is the extent to which the celebrity has been harmed by the manufacture and viewing (by one person only) of the deepfake$_{\text{private}}$. Certainly, she does not *experience* harm (whether physical or other), but this is the case with any deepfake$_{\text{p}}$ that the target of manipulation is unaware of. There is, however, an increased likelihood of her becoming aware of a deepfake$_{\text{p}}$—the longer it exists and the more accessible it is—and therefore of her experiencing harm as a consequence of this, including harm to her reputation. In principle, the same can be said about private deepfakes$_{\text{p}}$, of course, except, in the case of this particular private deepfake$_{\text{p}}$, no one had accessed and viewed it prior to its destruction other than myself, and, as it has now been destroyed, no one else will; neither will they gain knowledge of its fleeting existence, as I have not divulged (nor will I) what I did to anyone. As such, the celebrity will never come to experience harm or have their reputation damaged on account of the manufacture, brief existence, and limited viewing of the deepfake$_{\text{private}}$. Even the fact that *I* have knowledge of the deepfake$_{\text{private}}$ cannot damage the celebrity's reputation because, to reiterate, I will never divulge this fact to anyone and, given that I created the deepfake$_{\text{private}}$, I know it was a fiction.[7]

In doing what I have done, have I disrespected the celebrity? I believe so. As the deepfake$_{\text{private}}$ satisfies the conditions for something to be pornography—even with an audience of one—it remains an example of nonconsensual pornography. By creating nonconsensual pornography, I have acted disrespectful toward the celebrity (or whomever) because my action demonstrates a failure to regard them as someone who has intrinsic value in virtue of being a person (as discussed in chapter 11). Through the creation of the deepfake$_{\text{private}}$, irrespective of the fact that the target of manipulation is unaware of what I have done, and so does not experience harm on account of it, my intent to act was nevertheless such that, if the target of manipulation were to have discovered what I had done, and therefore what it erroneously *appears* that she has done, she would likely be hampered (to a greater or lesser degree) in the pursuit of a positive understanding of herself as a person with intrinsic value. In creating the deepfake$_{\text{private}}$, and in the absence of evidence to the contrary (such as coercion or naivety), not only was it my *intention* to do something that had the potential to impact negatively on another person, irrespective of whether it succeeded but, because of this, I demonstrated a willing disregard for the individual targeted.[8] As with the creation of the deepfake$_{\text{p}}$, I dismissed their *value* as a person (i.e., an autonomous, rational moral agent) with values of their own that (very) likely run contrary to what I did in creating the deepfake$_{\text{private}}$.

FANTASY

If creating a deepfake$_{\text{private}}$ under the conditions described is morally questionable, because it raises issues to do with consent and disrespect, then do these same issues apply to fantasy? To illustrate: suppose I am unaware of the existence of deepfakes$_{\text{p}}$ or, if aware, unable to access or create them myself. Instead, I fantasize about engaging in a sexual act with a celebrity (or whomever).[9] Morally speaking, is what I have done (i.e., engaged in a private fantasy about a nonconsensual sexual encounter) less morally problematic than producing my very own (private) deepfake$_{\text{p}}$ with essentially the same content, in terms of the consent required, likelihood of harm, and respect shown?

Fantasy and Consent

Recall from earlier discussion that idle fantasies have an autonomous quality and do not necessitate the existence of a corresponding desire for the content of the fantasy to be realized. The object of my fantasy has nevertheless not consented to be misrepresented, sexually, in this idle way, just as she would not have consented (I conjecture) to having her image inserted into my very own private deepfake$_{\text{p}}$. Given that consent is required in the latter case, on account of the pornographic nature of the deepfake$_{\text{private}}$, what, if anything, prevents its requirement in the former?

The lack of consent characteristic of *pornographic* deepfakes is sufficient grounds to object, morally, to their creation. Fantasies, however, no matter how sexually charged and graphic, are *not* a species of pornography: for definitions of pornography, while varied in other ways, tend to imply, or make explicit reference to, images or other materials (e.g., written or audio descriptions) used to depict the act in question (see Ashton et al., 2019, for a detailed overview). As such, it is illegitimate to level a charge of nonconsensual *pornography* at fantasies, even in the absence of consent, because the means by which fantasies are crafted does not satisfy the materiality implicitly or explicitly contained within definitions of pornography. Moreover, because of the absence of materiality, and therefore public accessibility, it is inappropriate to talk of, let alone require, consent from those who feature in fantasies, and this remains the case irrespective of content. It is one thing to say to someone, “I do not consent to you making available these images of me” and quite another to say, “I do not consent to you thinking about me.”

Deepfakes$_{\text{p}}$ are created using a medium that affords public accessibility. They are *disposed* to be accessible to others, all things being equal, even when the deepfake$_{\text{p}}$ is kept under lock and key or remains, in some other conventional way, private. Fantasies (as I am using the term), in contrast, are not produced using a medium that has the capacity to be made public directly; one’s *thoughts* are not disposed to do this, and therefore one’s exclusively thought-based fantasy is not disposed to do this, either.[10] This means that “private” is not an equivalent term when contrasting private deepfakes$_{\text{p}}$ with private fantasy.

To avoid a charge of equivocation, then, it is important to make clear that a private fantasy is not like a private diary entry or private conversation, as the latter two are publicly accessible (in principle) irrespective of agent disclosure, whereas the former is not (i.e., in the case of fantasy, public accessibility *requires* agent disclosure; although what is disclosed is, by its very nature, a mediated form of the fantasy). Given this, where the term “private” refers to that which is not disposed to be directly publicly accessible, consent should not be required. Of course, accepting this, and therefore judging that one’s fantasy does not violate any kind of consent requirement, does not immunize it against further moral scrutiny.

Fantasy and Harm

As with deepfakes$_{\text{private}}$, it is difficult to envisage how the object of fantasy would experience *harm* or have their reputation damaged simply as a consequence of someone fantasizing about them, even when the fantasy is sexually explicit and involves immorality (e.g., the aforementioned fictitious, nonconsensual, sexual encounter), unless, of course, the fantasy’s misrepresentation is made public: say, in the form of a deepfake$_{\text{p}}$. This medium’s potential for harm has already been discussed in chapter 11, however.[11] Before considering the harm of fantasy, specifically, then, it is worth restating that I am talking about idle and not surrogate fantasy. With the former, I do not desire to realize the content of the fantasy or hold beliefs congruent with it. Therefore, I do not *harm* the object of the fantasy by perceiving them as someone toward whom or with whom it is acceptable to act immorally (in ways other than the presently debated act of fantasizing about them, of course). (Recall from chapter 5 that endorsing an immoral worldview about the acceptability of sexual violence toward women, for example, makes things worse for, and therefore harms, women because, at the very least, they are *regarded* as persons toward whom violence is acceptable, even if one never inflicts such violence on them, or attempts to do so.) That said, by featuring them, even in an idle fantasy, in the absence of any form of coercion (broadly construed), at the very least, I must desire them to feature in the fantasy (to fantasize about them) and, quite possibly, hold the belief that it is acceptable to treat them in accordance with this desire (i.e., as the object of a sexual fantasy in which they are misrepresented as engaged in nonconsensual sex). Where this is the case, is such treatment of another person morally acceptable or does it demonstrate a lack of respect? To be clear, the putative disrespect is not because I regard the object of my fantasy as someone against whom it is acceptable to engage in sexual violence, because I do not; it is an idle not a surrogate fantasy. In the next section, I explore this question further.

Fantasy and Respect

Suppose I fantasize about engaging in nonconsensual sex with the celebrity featured in the magazine I have just been reading, or my new neighbor, or the stranger I have just noticed sitting opposite me in the café. Whomever I choose as the object of my fantasy, they have not consented to be used as such; but, as discussed, consent should not be a moral requirement as far as fantasy is concerned. Neither do they *experience* harm as a result of my *idle* fantasy (again, for reasons discussed). Despite this, is what I am doing disrespectful?

Through my act of fantasizing, I am prepared to misrepresent sexually—even if only to an audience of one (namely, myself)—the object of my fantasy: by fictively imagining a counterfactual condition in which (in this case) the stranger sat opposite me in the café, contrary to reality, is engaged in a nonconsensual sexual act with me. Is such a fantasy disrespecting the object of the fantasy, irrespective of whether they would *feel* degraded and disrespected by it, should they ever find out? After all, it is not inconceivable that they would become excited at the thought of "starring" in my rape fantasy, or perhaps have no strong feelings either way.[12] Nevertheless, where the wishes and/or values of the object of my fantasy (whatever they happen to be) have been dismissed, then I am treating that person disrespectfully. I am treating them *only* as a means to an end (to draw from Kant, [1785] 1993)—that is, only as a means of securing my own pleasure—and not as an end in themselves. To be clear, I do not mean that I am treating them only as a means to an end simply because I fantasize about them irrespective of their wishes but, rather, because in order to do this, I must be of the opinion—or hold the belief—that they are someone toward whom I can act in this way.

In contrast, suppose I fantasized about living in a mansion with my family and never wanting for anything. Does such a fantasy disrespect them? It misrepresents our current situation and therefore misrepresents them to some degree. I am also using my family as a means to an end: to create a pleasant fantasy. But do I treat them *only* as a means to an end? No, because I have reason to believe that such a fantasy, even if not shared, would not run contrary to their values. I therefore have reason to believe that they would not begrudge me this fantasy. By taking into account their standing as persons and doing something that either aligns with their values or at least is not an affront to them, I am not *treating* them simply as a means to an end but an end in themselves, even if it turns out that I am mistaken about their views on this particular fantasy. I am therefore confident that I am able to perform this private act without disrespecting them.

Suppose, then, I ask myself: Would the stranger in the café object to my idle (not surrogate) fantasy about the nonconsensual sexual encounter (attack)? To be clear, I am not suggesting that one should have to deliberate prior to the commencement of a fantasy. It may be that the thought/image of the sexual assault suddenly pops into one's head, as an unbidden thought over which one has little control. Under such a circumstance, I am suggesting instead that one has a choice either to disregard the thought or to entertain and embellish it. It is at this point that the aforementioned question should be considered.

Asking the question requires that I recognize the stranger's standing as a person. Suppose I decide that there is a good chance that she would object, but I continue with the fantasy anyway. In so doing, I have chosen to disregard what I believe are the likely values of the stranger, thereby disrespecting her. However, suppose I sincerely believe that the fantasy's content would not run contrary to her values and that she would therefore not object. Even if I am mistaken, I have sought to take into account her values, at least as I believe them to be, and so cannot be said to have acted out of disrespect.

A point of clarification is required. I would need to have a good (*qua* justified) reason to believe that the fantasy would not go against her values. In most cases, this is a requirement I am unlikely to be able to satisfy. Moreover, the reason cannot be based on some demeaning stereotypical view of her, as this itself would be disrespectful.[13] Owing, therefore, to the difficulty in justifying the claim that a stranger, or a celebrity I do not know, would not object to being the object of my fantasy—especially of an immoral sexual nature—there is a strong likelihood that to continue with the fantasy would require that I dismiss whatever unknown values the person holds out of hand and, in doing so, disrespect them. Importantly, though, unlike the creation of a deepfake$_p$, which would typically require the same dismissal of their values, my fantasy does not constitute the *public expression* of my disrespect for the person targeted. One's fantasy is not disposed to be publicly accessible directly. Therefore, one fantasy-based disrespect is not disposed to be directly publicly accessible.

Should this difference matter? After all, according to Kant and Eidelson, we are morally obliged to respect persons. Perhaps, instead, the moral wrongness of the fantasy-based disrespect admits of degrees, whereby the lack of public expression tempers the wrongness, relegating it to what Smuts (2016) calls *disesteem* and Driver (1992) describes as a *suberogatory* wrong. Notably, something that is not forbidden but which is morally frowned on. Or if immoral, it is nevertheless not the most severe immoral act one could engage in. Degrees of wrongness regarding one's disrespect and its public expression strike an intuitive chord. It seems right that the disrespect shown to the object of one's fantasy is less morally objectionable than the disrespect shown to the person whose face is superimposed onto a deepfake$_{\text{private}}$, which, in turn, is less morally objectionable than the disrespect expressed through the publicly available deepfake$_{\text{p}}$. This is especially so when the fantasy is idle and therefore does not involve the belief that the subject of one's fantasy is the sort of person who one would like to rape for real (for example) but, rather, merely fantasize about in this way. On the other hand, it may be that the disrespect is equivalent in each case but, with pornographic deepfakes and surrogate fantasy, the wrongness is compounded by other factors (as discussed), thereby reducing idle fantasy to a less severe form of wrongdoing, all things considered.

CONCLUSION

In this chapter I have presented a nuanced account of the morality of fantasy, by contrasting idle fantasy with surrogate fantasy. I have argued that engaging in surrogate fantasy is immoral because the surrogate fantasy is borne of immoral beliefs and/or desires. And while this is not the case with idle fantasy, or at least not to the same extent (because of its autonomous nature), idle fantasy is still not immune to moral scrutiny, particularly where the act of fantasizing is considered disrespectful, owing to a lack of regard for what the objects of the fantasy's views are on their role in the fantasy. In comparison to pornographic deepfakes, however (even private ones), idle fantasy's lack of capacity for the direct public expression of disrespect potentially mitigates its wrongness because there is no cumulative wrongness to consider. That is, no additional concern over a lack of consent and/or the risk of harm that might compound the wrongdoing. I will, however, return to this matter in the next chapter when discussing constructive ecumenical expressivism and the construction of moral norms (see the section "The Compatibility of the Moral Claims Made About Fictional-x_i with CEE"). This discussion forms part of a broader objective: namely, differentiating between poor taste and immorality in order to provide a more nuanced measure of the wrongness of depictions of fictional immorality.

NOTES

1. https://www.telegraph.co.uk/men/thinking-man/terry-waite-i-spent-five-years-as-a-hostage-in-beirut---but-i-ne/ (accessed August 22, 2019).

2. The only way the world cannot be is identical to my fantasy, as such an eventuality would prevent it (the fantasy) from being a counterfactual event.

3. By "voluntarily control," I mean to discount extreme examples of exerting control, such as allowing oneself to be put into an induced coma.

4. Meaning "for want of a better alternative."

5. Recall that the *clinical* use of the term "pedophile" is reserved for those who have a sexual interest in prepubescent children but, also, that the term is often adopted by the media and used in common parlance to refer to those with a sexual interest in minors of any age. Recall, also, that I have adopted this more common usage throughout the monograph, unless otherwise stated, as in the case here, where the research discussed has tended to focus on prepubescent children.

6. Biovana and Critelli (2009) conceive of rape fantasy along an erotic-aversion continuum with aversive fantasy lacking in erotic content or failing to elicit pleasurable feelings.

7. This particular deepfake$_{\mathrm{p}}$ satisfies the criteria I have been using for fiction (as I am both the creator and the intended audience). It is, therefore, correct to refer to it as a fiction rather than a fake.

8. I say this while assuming the absence of pathology, such as the delusion belief that the celebrity would share the desire to appear in the deepfake$_{\mathrm{p}}$.

9. As an aside, I recognize, of course, that watching a deepfake$_{\mathrm{p}}$, or in fact engaging with any other fiction (or nonfiction, for that matter) may be a way of facilitating my fantasy, but need not be; I can simply close my eyes and fantasize.

10. Of course, with the advent of capable technology, should it become the case that one's thoughts are publicly accessible, in a manner equivalent to the public accessibility of one's private diary, then one would have to rethink this position.

11. Again, I use the term "fiction" (in inverted commas) because it may be that the fiction is presented in such a way as to not satisfy the criteria of fiction I am using (see chapter 3).

12. Imagine, for example, that one revealed to one's sexual partner that one fantasized about engaging in nonconsensual sex with them. On hearing this, it is not inconceivable that the partner would be excited by the fantasy. To be clear, in offering this as a possible response, I assert only that it is not inconceivable, not that it is typical. In fact, against the likelihood of this, see Seuntjens' (2013) article on the *masturbation fantasy paradox*: the idea that people typically do not engage (or find it difficult to engage) in a masturbatory fantasy that features the person they are in love with.

13. In the case of someone with a cognitive impairment, and therefore constrained rational autonomy and ability to establish their own values, I should base my judgment on the established social values of the community to which I belong (killmister, 2017), and not on what I take the person's cognitively impaired values to be.

Chapter 13

"It's Not Immoral, but It Is in Poor Taste"

INTRODUCTION

In chapter 9, I discussed what it is for something to be in poor (moral) taste, such that, for *x* to be in poor taste, necessarily, S perceives it to be offensive. Moreover, the reason S deems *x* to be offensive is that she believes it realizes some property (O) of which she disapproves (*qua* is offended by) in virtue of *x*'s treatment of a morally pertinent matter (e.g., trivializing murder or sexual assault, and so on). In light of this description, imagine you are watching television with a friend: enjoying a comedy sketch, perhaps, or maybe a documentary or an interview, or simply watching an advertisement. Or perhaps, instead, the two of you are playing a video game. Whatever the circumstance, at some point, your friend turns to you and, with reference to something said or done on the screen declares, "*That's* immoral." After some thought, you respond: "It's not immoral, but it is in poor taste."[1]

You and your friend disagree over the moral status of *x*, insofar as your friend believes it to be immoral and you do not; yet you add that *x* is in poor taste. How should we understand the nature of this concession? You did not assert that *x* is simply a *matter of taste*; rather, you stated that it was in *poor* taste, which suggests a standard of taste against which *x* should be judged and, in this case, found wanting. Whatever this standard, it needs to be in some way distinct from your measure of morality, given that you claim that *x* is not immoral and presumably are not trying to contradict yourself.

In this chapter, I intend to show that pronouncements about immorality are a product of the same underlying process as pronouncements about poor taste, insofar as each conveys a negative attitude toward *x* based on a belief that *x* realizes a property of which S disapproves. Given this, to prevent the proposition "*x* is not immoral but is in poor taste" from being morally contradictory, because it appears to be intimating that *x* is both *not* "something that ought not to done" (on account of *not* being immoral) but *is* "something that ought not to be done" (on account of its poor taste), it is incumbent on me to identify a morally relevant means of discriminating between poor taste and immorality based on their respective reasons for intimating behavioral prohibition. This move is important if we are to uphold the validity of degrees of wrongness regarding depictions of fictional immorality, such that pronouncing that *x* is not immoral but is in poor taste picks out a morally coherent distinction. It is toward a defense of this position that I now turn.

THE PROBLEM OF CONTRADICTION

Prima facie, the proposition "*x* is not immoral but is in poor taste" (hereafter, NI_1) is not contradictory. Nevertheless, recall from chapter 9 that:

(a) Pronouncing that "ϕ is immoral" intimates that ϕ ought not to be done.
(b) Pronouncing that "ϕ is in poor taste" intimates that ϕ ought not to be done.

Nominally, (a) and (b) intimate the same behavioral outcome: namely, not doing ϕ. When NI_1 is reconfigured to show this behavioral similarity, it reveals a contradiction, as NI_2 illustrates:

> NI_2: *x* is not "something that ought not to be done" but is "something that ought not to be done."

The contradiction can be easily avoided, of course, if one takes into account the *reason* one ought not to do *x* in each case, and this reason is sufficiently different. In the case of immorality, let us identify the reason (for now) using the demonstrative pronoun "this," and, in the case of poor taste, using the demonstrative pronoun "that," so that we get:

> NI_3: *x* is not "something that ought not to be done for *this* reason" but is "something that ought not to be done for *that* reason."

In addition, one might seek to avoid the contradiction by appeal to equivocation: insofar as, in both NI_2 and NI_3, the use of "ought" is ambiguous and potentially plays a different role each time it is used within the proposition (recall discussion on the different uses of "ought" in chapter 9). Reference to the reason *x* ought not to be done is therefore likely to provide some indication of which "ought" is being employed, as NI_3* and the accompanying example illustrate:

> NI_3*: *x* is not "something that ought* not to be done for *this* reason" but is "something that ought** not to be done for *that* reason."

Thus, *x* (*qua* buying chocolates as a leaving gift for S) is not "something that ought* not to be done because it is unkind" but is "something that ought** not to be done because S does not like chocolate." Here, "ought*" is indicative of what, for now, I will refer to as a moral ought, whereas "ought**" indicates a nonmoral or practical ought. When one differentiates between the behavioral outcomes intimated by the respective claims of immorality and poor taste, in virtue of the occurrence of a different *reason* for their prohibition, and likely a different "ought," the contradiction evident in NI_2 dissolves.

Where *x* is immoral, in a sense, the *reason* one ought not to do *x* has already been provided: namely, because it is immoral. *Mutatis mutandis*, the same can be said about poor taste. Of course, what is missing from either of these reason-giving explanations is any reference to why being *immoral* or in *poor taste* "ought" (in a manner yet to be more clearly defined) to be prohibited. This further reason is important to establish: for not only does it account for why we should not engage in *x* beyond the circular assertion that we ought not to engage in the immoral because it is immoral, or poor taste because it is in poor taste; more than this, it provides the means of differentiating between the two, at least *if* the reason for prohibition can be shown to be sufficiently different. Should there be no legitimate reason to differentiate between immorality and poor taste, then the prima facie noncontradictory conjunction presented in NI_1 would be undermined owing to the invalidation of the putatively different reasons for prohibition expressed within NI_3/NI_3*, which serves to rescue NI_1 from the explicit contradiction found in NI_2.

For the moral realist, the distinction between taste and morality is straightforward enough. Where the moral realist believes that *x* is immoral, she likewise believes that *x* violates some mind-independent principle. Where *x* is said not to be immoral but, rather, in poor taste, *x* is accused of violating some other, non-mind-independent, principle. But what if one is not a moral realist and, instead, considers moral pronouncements to be indicative of one's attitude? From an expressivist's perspective (although I intend to narrow the focus of discussion to a particular version of expressivism, shorty), whereby pronouncements about taste and morality convey one's attitude, rather than pick out some mind-independent aspect of reality, how are we able to differentiate between immorality and poor taste? In other words, what makes one's negative attitude toward x_1 congruent with one's declaration that x_1 is immoral, but in the case of x_2, *mutatis mutandis*, congruent with one's complaint about taste?

In rejecting moral realism (without defense), I intend, instead, to champion a version of expressivism known as *constructive ecumenical expressivism* (CEE) (Young, 2014b, 2015).[2] As we shall see, however, CEE faces a problem insofar as it holds that judgments about immorality are arrived at through the same process as judgments about poor taste (see chapter 9). Yet, the utterance "*x* is not immoral but is in poor taste" does not *appear* bizarre, incoherent, or morally contradictory; rather, it is typically understood to be picking out something of note: a wrongdoing of some sort, even if it is not an immoral one.[3] NI_1 therefore has intuitive appeal. Given this, the task I have set myself is to extol the explanatory virtues of CEE while upholding the noncontradictory status of NI_1. To succeed, I have to show that CEE has the resources to provide a morally relevant means of differentiating between poor taste and immorality and therefore provide a relevant means of distinguishing between judgments arrived at through the *same* underlying process of attitude formation that intimate the *same* nominal behavioral outcome (namely, not doing *x*), but which are not morally contradictory. In short, I intend to show that:

(1) Judgments about poor taste and immorality are the product of the same underlying process.
(2) Given (1), each amounts to the expression of an attitude toward *x* based on a belief about a property *x* realizes and, in the case of *poor* taste and *im*morality, one's disapproval of this property.

Moreover, in line with the argument set out in chapter 9 in which:

(3) Judgments about poor taste reflect one's negative attitude toward the *treatment* of a morally pertinent matter; that is, a comment is said to be in poor taste *if* it is believed to *trivialize* something one considers to be immoral (inter alia racism, sexism, sexual assault, torture) or moral (e.g., charitable fundraising).[4]

I intend to show that:

(4) Given (3), poor taste is parasitic on the morally pertinent.
(5) Given (3), NI_1 is not morally contradictory from the perspective of CEE.

CONSTRUCTIVE ECUMENICAL EXPRESSIVISM

Where S declares that *x* is immoral:

(CEE) S disapproves of P and believes that *x* realizes P.

Property (P) can and does amount to different things for different people (subsumed under property, P, is property *p* or *q* or *r* or *s*, and so on). For S_1, P may amount to negative utility—the realizing of more displeasure than pleasure in the form of increased harm—while S_2 may hold it to be a violation of God's law or constitutive of a failure in one's secular duty to others. S_3, in turn, may characterize P as a vice rather than a virtue, and so on. Under CEE, declaring that "*x* is immoral"—where *x* equates to murder—allows (inter alia) the following possibilities:

(CEEa) A disapproves of *p* (where *p* equates to a violation of God's law) and believes that *x* realizes *p*.
(CEEb) B disapproves of *q* (where *q* equates to increased unpleasantness) and believes that *x* realizes *q*.

What CEE allows, as illustrated through CEEa and CEEb, is that A and B have a shared negative attitude toward *x*, which they expressed through their shared declaration that "*x* is immoral." In effect, A and B have the same *de re* attitude toward *x*. But this shared *de re* attitude is the result of the belief that *x* realizes some property (P) that they both disapprove of but which can be (and is) different for A and B. A believes that *x* realizes *p* and B believes it realizes *q*. Their differing belief about which property is realized by *x* means that they have different reasons for their shared *de re* attitude. In effect, they have different *de dicto* attitudes: that is, different beliefs about why it is immoral.

When S declares that *x* is immoral, one might be forgiven for thinking, in traditional expressivist style, that S is simply *expressing* a negative attitude toward *x*—something like "Boo *x*!"—which cannot, of course, be truth-apt (Jackson & Pettit, 1998); but, in fact, while the proposition "*x* is immoral" is congruent with S's negative attitude toward *x*, it should not be thought of simply as a means of expressing this attitude. Instead, through the proposition "*x* is immoral," S is *reporting* her belief that *x* is immoral in virtue of realizing P, a property she happens to disapprove of. Thought about in this way, what S has to say is truth-apt. Given this, where a shared (*de re*) attitude occurs with regard to an object or event, as a society we are able to create or *construct* a social norm that acquires its own intersubjective moral standard (Prinz, 2007). As McAteer (2016) explains when drawing on the philosophy of David Hume:

> To call something an intersubjective reality is to distinguish it both from objective and subjective reality. Something is *objective* if it is mind-independent, i.e., if it exists independently of all mental representation. Something is *subjective* if it is individually mind-dependent, i.e., if it exists only in one person's experience and is hence relative to that person's individual point of view. Something is *intersubjective* if it is collectively mind-dependent, i.e., if it exists in a group of people's experience such that it is relative to what Hume [calls] a "common" or "general" point of view. (McAteer, 2016, p. 14)

It is important to note that McAteer talks about a common point of view with reference to a shared way of *experiencing* something. I wish to talk, instead, about a common point of view in attitudinal terms, something endorsed by CEE.

The moral norm that emerges through the force of social consensus (*qua* a shared *de re* attitude) creates a common point of view that amounts to "a kind of objectivity in that it is not relative to any individual person's thoughts, feelings, or desires" (McAteer, 2016, p. 16). This kind of objectivity provides a normative standard against which individual actions are morally scrutinized.[5] As such, when S shares the same attitude toward the immoral as her society, she will be commended for doing so, even if only tacitly: for her attitude accords with the constructed moral norm's intersubjective normative status. When S does not, her society will feel it appropriate to rebuke her for her alternate, some might even say deviant, moral attitude. They will feel it appropriate to do so because both the rebuke and a change of attitude on the part of S are (believed to be) warranted (Nichols, 2008).

What A and B have in common is their negative (*de re*) attitude toward *x*. However, this singular attitude is adopted by A and B for different reasons (they have different *de dicto* attitudes). It is therefore my contention that a *de re* attitude shared by the majority of people within a given society (such that it becomes the constructed moral norm of that society) is more robust if it is the product of a number of different *de dicto* attitudes (i.e., if it is based on a number of different reasons for having the moral attitude). This should not be taken as evidence of inconsistency, and therefore as a reason to undermine the normative authority of the moral attitude; rather, and to reiterate, it should be taken as evidence of its robustness, insofar as there are purportedly many reasons why *x* (whatever *x* happens to be) is immoral. It just so happens that different people have different views on what these reasons are or how to prioritize them.[6] To undermine the moral (*de re*) attitude, one would have to undermine the various reasons (*de dicto* attitudes) justifying its intersubjective normative status. Such a requirement does not rule out a change of *de re* attitude—there is therefore a degree of fluidity inherent within CEE—but it does make any change less capricious.

One might think of the objectified nature of the intersubjective moral norm as a kind of socially evolved version of Hume's expert critic (Hume, [1757] 1995). The role played by the critic in Hume's philosophy is to establish a standard for *aesthetic* taste against which one's personal taste could (should) be measured. Similarly, the constructed social norm I am proposing determines the appropriateness of one's moral attitude. To reiterate, the robustness of the normativity measure is supported through the number of disparate reasons for arriving at one's moral attitude. Under CEE, *x* is immoral not because it possesses some independent immoral property but, rather, because of a given society's shared attitude toward *x* that is itself, in the more robust cases, borne of a number of disparate reasons for *x*'s immorality. In short, it is our shared attitude toward *x* (as immoral) that makes *x* immoral.

THE PROBLEM OF MORAL ERROR AND CONFLICTING NORMS

If moral judgments are attitudinal and the "correct" judgment is a construct, then, while it may be that individual judgments can be deemed correct or incorrect, or appropriate or inappropriate, relative to the intersubjective standard, how can the intersubjective standard itself ever be in error or be considered inappropriate? The intersubjective standard is a product of the disparate reasons that culminate in its formation. Each of these is constitutive of a belief (or series of beliefs) that can be scrutinized. To illustrate, and borrowing from Blackburn (2009), a priori, if S believes that "F is A," and such a belief is able to withstand all attempts to improve it, then the belief must be true. CEE does not demand that the intersubjective attitude is based on beliefs resistant to improvement. Suppose, instead (and, again, following Blackburn), that S's belief is resistant to anything S would *recognize* (and therefore accept) as improving her belief. Given this, it does not follow that S's belief is true. Nevertheless, S is likely to feel secure in her belief that "F is A" but, importantly (*and* unfortunately), not be in a position to be made aware of any improvement that, say, the belief "F is *sometimes* not A" would make to her moral outlook (her attitude).

There is, however, an important difference between being *unable* to recognize improvements to one's belief and being *unwilling* to recognize them. I take as a given the fact that all of our divergent societies are sufficiently similar, cognitively and conatively speaking, to be capable of the same perceptual and reasoning skills and therefore capable of recognizing improvements to beliefs. In accordance with this view, CEE demands that S is secure in her belief that "F is A" not because she is unable to recognize improvements to her belief but because she has not *yet* been exposed to a means of improving her belief (in a way that she is able to recognize, regardless of her willingness to do so). What this allows is security of belief in the absence of occurrent awareness of the need for belief revision but, importantly, the *possibility* of belief revision and so attitudinal change (should a change in belief warrant a change in attitude). What it does not require is that S's belief be true; nor does it set truth as the ultimate goal for S's belief.

Given this, it is possible, and in fact historically and presently the case, that different *de re* attitudes exist across different cultures/societies (e.g., attitudes toward homosexuality or premarital sex, or, more historically, slavery). While accepting relativism, insofar as our moral position is relative to a particular intersubjective norm, it does not follow from this that we should always tolerate differences constitutive of different intersubjective norms, and therefore never be justified in privileging one attitude (whether *de re* or *de dicto*) over another (see below). CEE therefore provides an antirealist means of justifying the claim that a particular society's attitude to *x is* more appropriate than another's—meaning it should be privileged over another's—without having to endorse any form of realism, quasi or otherwise. By this, I mean without having to succumb to the quasi-realist's agenda of trying to reconcile the "expressive" or attitudinal ontological status of morality with realist demands for moral natural kinds. It also allows "common sense" morality to be explained from a relativist perspective. It explains why, for example, one should privilege a moral position that advocates equality of gender over one that promotes the subjugation of woman, or freedom for all over selective slavery, or prohibits nonconsensual sex within marriage. It also accounts for why the view that one should not set cats' tails on fire for fun should be privileged over one that states we should, or that expresses indifference to the plight of cats; and it does this without having to fall back on the existence of natural moral properties of the world.

The fact that CEE is a form of moral relativism is therefore not something I take to be a weakness of the approach. This is because CEE incorporates a form of moral relativism that enables one moral (*de re*) attitude to be privileged over another whenever this attitude is more robustly constructed. The more robust attitude *should* (in a rational sense) be privileged because what makes it more robust is its closer alignment with available evidence and reasoned argument (in virtue of at least some of its many reasons for the attitude). An attitude built on different beliefs (*de dicto* attitudes), where a number of these are presently secure but open to the possibility of revision (say based on updated empirical evidence or the identification of biases, or prejudices or other flaws affecting one's reasoning), proffers an intersubjective standard that more securely grounds the (*de re*) attitude. This means that within a group of disparate beliefs supporting a particular negative attitude toward *x* (e.g., pedophilia), if one (or possibly more) of these beliefs (*de dicto* attitudes) is shown to be erroneous, then the *de re* attitude is likely to remain intact given that there are other *de dicto* attitudes supporting it. If, for example, S has a negative attitude toward pedophilia *only* because she believes it necessarily violates either fidelity within marriage or celibacy outside of it, then the basis for her attitude can be shown to be problematic without necessarily undermining any of the other reasons provided for maintaining a negative attitude toward pedophilia.

Robustness does not prevent a change to a *de re* attitude from occurring, of course. To reiterate, CEE allows for change: there is a certain fluidity built into the position should enough *de dicto* attitudes be revised and should these revisions necessitate a change in the *de re* attitude. But it does make change less capricious, owing to the fact that it is not necessarily dependent on any particular idiosyncratic and/or entrenched belief.

The Compatibility of the Moral Claims Made about Fictional-x_i with CEE

The aim of this book is to identify necessary and sufficient conditions for the immorality of depictions of fictional immorality. At first glance, there appears to be a tension between these conditions and CEE: for none makes reference to social consensus or expressions of negative attitude, or any characteristic of CEE. This apparent tension is mistaken, however, as I shall now demonstrate using the example of IWV. Recall that the first part of IWV reads:

- *Where a fiction, taken as a whole, endorses an immoral worldview, then . . .*

This clause has nothing to say about *how* the immoral status of the worldview the fiction is (allegedly) endorsing is determined. A version of the clause that makes explicit reference to CEE would look like this: where a fiction, taken as a whole, endorses an immoral worldview that is deemed immoral in virtue of a robust shared *de re* attitude, then . . .

- . . . *any depiction of immorality from within the narrative, congruent with that worldview, is immoral.*

The original second clause of IWV (above) applies whenever the following conditions are met: the worldview depicted is immoral and fictional-x_i is endorsing that worldview. Given this, *reason* dictates that if, in accordance with CEE, I have a negative attitude toward a particular worldview—say, white supremacy—because I believe that it realizes a property of which I disapprove, then I will have a negative attitude toward that which endorses this worldview (given that it is endorsing something of which I disapprove). In short, if I believe that A is immoral and that B endorses A, then it seems reasonable to surmise that I will believe that B's endorsement of A is immoral. Because of this, IWV is not in conflict with CEE's fundamental assertion that immorality is attitudinal, nor that its (CEE's) normative credentials are established through robust social consensus. For where the normative position arrived at through the application of CEE is that x is immoral (and therefore ought not to be done), and where fictional-x_i endorses x, then it follows that fictional-x_i, in virtue of endorsing x, is immoral (and ought not to be done), just as IWV prescribes. Or, put differently, where a robust moral norm, established through social consensus, holds that x is immoral, reason dictates that a fictional depiction of x that falls foul of IWV would, in virtue of this moral norm and of violating IWV, be likewise deemed immoral. While the relationship I am describing does not constitute a logical entailment, it does establish coherence.

The same move can be made with the other conditions for the immorality of fictional immorality identified. For the sake of brevity, however, I will support the assertion that each is compatible with CEE with two further examples only. First: If, in accordance with CEE, one holds that x is immoral, then, given what it means to enjoy$_{(sub)}$ fictional-x_i, it follows, insofar as coherence of belief dictates, that one would hold that enjoying$_{(sub)}$ depictions of fictional-x_i is immoral. Indeed, it is because CEE does not consist merely of attitudinal expressions—such as "Boo x!"—but also beliefs about properties realized by x that one expects—requires, even—relations between beliefs about x, including one's attitude toward x borne of these beliefs, to be consistent and coherent. If I believe that x is immoral in virtue of realizing a property of which I disapprove, then a rational network of beliefs should lead me to conclude that delighting in the idea of that of which I disapprove, or desiring to actualize that of which I disapprove (both of which constitute a motivation for enjoying$_{(sub)}$ fictional-x_i), is similarly immoral.

Second: Recall how I argued in the previous chapter that idly fantasizing about sexually assaulting a stranger (in the café) is disrespectful, but also that the lack of public expression accompanying this act of disrespect is a mitigating factor that relegates *this* type of disrespect to a lesser wrong. It is a lesser wrong that can nevertheless be elevated if the disrespect is revealed. This is because disrespect that is revealed could potentially harm the target of the fantasy, thereby compounding the wrong, as would likely become apparent when witnessing the individual's reaction to the revelation. In other words, even though the belief I hold about the person in my idle fantasy is disrespectful, insofar as it is the belief that they are someone whose (potential) views on whether they want to be the object of *my* fantasy can be disregarded (a disrespectful way of thinking about anyone), it is still less disrespectful than the belief I would necessarily hold toward them in the case of a surrogate fantasy involving sexual assault, or if I were to feature them in a pornographic deepfake.

In accordance with CEE, then, the extent to which the wrongness of idle fantasy is comparable to the wrongness of surrogate fantasy or public or private pornographic deepfakes (namely, whether they are immoral, immoral but less severe, or an altogether lesser wrong) will depend on the degree of moral consensus that exists on the matter. The degree of consensus is synonymous with the similarity of the *de re* attitude toward the type of idle fantasy described (i.e., that it is something people find immoral in virtue of their belief that the fantasy realizes a property they disapprove of) and how this compares to their disapproval of the property they believe is realized by surrogate fantasies and public and private pornographic deepfakes. I am of the opinion that there is currently no consensus on the nature or strength of the wrongness of the kinds of idle fantasies we have been discussing.

A POTENTIAL PROBLEM: DISTINGUISHING BETWEEN P AND O

Returning to the issue of taste and immorality, as we have seen, where one declares that "*x* is in poor taste," one is in effect declaring (necessarily) that one has a negative attitude toward *x*: that *x* is offensive in virtue of one's belief that it realizes property O (of which one happens to disapprove). Of significance, then, but also something that is potentially troubling if one wishes to differentiate between immorality and poor taste from the perspective of CEE, is this: according to CEE, where one declares either that *x* is in poor taste or that *x* is immoral, the same underlying process is involved in assigning a negative attitude to *x*, as NI_4 illustrates:

> NI_4: *x* is not "something that ought not to be done because one disapproves of P and believes that *x* realizes P" but is "something that ought not to be done because one disapproves of O, and believes that *x* realizes O."

While NI_4 is not logically contradictory, what remains unclear is why disapproving of P, particularly given that P can be a different property to different people, is sufficient to make *x* immoral in virtue of believing that *x* realizes P, whereas, in the absence of the belief that *x* realizes P, the presence of the alternate belief that *x* realizes O, where one disapproves of O, does not make *x* immoral but, instead, an example of poor taste. Given that O, like P, can refer to different things to different people—for example, ridiculing or trivializing—O is no different to P as far as permitting different token reasons (*de dicto* attitudes) is concerned. Because of this, what still eludes us is an understanding of what it is about P, albeit in the form of different tokens of P (*p*, *q*, *r*, and so on), that makes disapproving of P sufficient to assign a charge of *immorality* to that which one believes realizes P—namely, *x*—compared to when one disapproves of *x* in virtue of believing that *x* realizes O, albeit in the form of different tokens of O.[7] This understanding is important if NI_4 is to avoid the charge of being *morally* contradictory.

Is the difference essentially that poor taste denotes one's negative attitude toward an act that is deemed to be wrong but not *that wrong* or bad but not that *bad*, at least compared to an *immoral* act? There is some truth to this (recall the discussion on suberogatory wrongness in chapter 9): for it is hardly remarkable to note that being guilty of trivializing murder is not as bad as committing murder.[8] But moral demarcation requires more than simply delineating degrees of wrongness constitutive of different levels of disapproval. After all, it is generally recognized that murder is worse than assault (say, when using the metric "level of harm inflicted"), but both are nevertheless *immoral* acts. One would be unwilling, I suspect, to relegate assault to the status of poor taste simply because it is not considered to be as bad (i.e., causes less harm) as murder. Is there, then, a difference in kind between disapproval in cases of immorality compared to poor taste that is able to explain their different categorizations? Put another way, while accepting differences in degree between immoral acts (e.g., murder compared to assault), and likewise between examples of poor taste (arguably, the portrait of convicted child murderer Myra Hindley created from children's hand prints compared, say, to the advertisement for liquid soap depicting a blood-stained arm next to a murder victim, and the caption "when ordinary soap just won't do"), are there nevertheless things common to poor taste that are absent from cases of immorality?

What tokens of O have in common—which identifies them as tokens of O and therefore as offenses characteristic of poor taste—is how they are construed as *treating* actions already identified within a society as morally pertinent. To illustrate, consider the frequently mentioned advertisement for liquid soap depicting a (fictitious) murderer's blood-stained arm beside his victim. For S to be offended by *x* (the advertisement), and therefore consider *x* to be in poor taste (hereafter x_{pt}), S must believe that x_{pt} is treating something that S already finds morally pertinent—namely, murder (x_m)—in a way that S finds offensive. Suppose, then, S_1 finds x_m immoral because it violates God's law (although it may equally be because S_1 believes x_m amounts to a failure in one's secular duty to others or because it increases harm, or is a vice, and so on). In accordance with CEE, we would say that S_1 believes that x_m realizes some property (P) that, in this case, amounts to a violation of God's law. Let us call this token of P, *vgl*. S_1 therefore finds x_m immoral in virtue of disapproving of *vgl* and believes x_m realizes *vgl*. For S_1, what makes *vgl* sufficient justification for a claim that x_m is *immoral* is the fact that S *believes* that God's law is a measure of *morality*, and therefore it follows from this that violating God's law is an immoral thing to do. Where one finds x_m immoral in virtue of *vgl*, then one is claiming that x_m is immoral because of what one believes about *vgl*. S_2, however, does not believe in God's law but does believe that murder is immoral because it amounts to a failure in one secular duty to treat a fellow human being as an end in themselves. Let us call this token of P—*qua* a violation of one's secular duty—*vsd*. S_2's negative attitude toward x_m therefore amounts to the same expression of immorality as that expressed by S_1, only, this time, it is because S_2 believes that x_m realizes *vsd* and *vsd* is believed by S_2 to be *immoral*.

As already noted, according to CEE, S_1 and S_2 have a shared *de re* attitude toward murder. They believe it is immoral, only for different reasons: reasons (or *de dicto* attitudes) that stem from their disparate beliefs about what counts as a measure of *morality*.[9] Should S_1 and S_2 also hold that x_{pt} (the advertisement) is in poor taste, then, according to CEE, this is because they believe that x_{pt} realizes O (*qua* some respective token of O). In other words, they each believe that the manner in which murder (which they both hold to be immoral, only for different reasons) is being treated by x_{pt} (e.g., in a trivializing way) is offensive. S_1 and S_2's shared negative attitude toward x_{pt} does not amount to an expression of immorality because they do not believe that trivializing murder is a violation of God's law or a failure in one's secular duty (for example); rather, they believe it is an expression of poor taste for the reasons discussed. Alternatively, it may be that S_2 does not believe that x_{pt} realizes O. In which case, despite S_1 and S_2's shared view that murder is immoral and therefore that x_m is immoral, they disagree over whether x_{pt} is in poor taste as a result of its treatment (as they see it) of murder. Unlike S_1, who finds x_{pt} offensive, S_2 finds it mildly amusing, rather than offensive (let us allow). Certainly, she does not construe it as trivializing or otherwise acting disrespectfully toward the victims of murder.[10]

Recall, also, from chapter 9, that the question of whether an act of *murder* has actually occurred is less ambiguous than whether one is treating murder in a trivial way. It is therefore much more likely that a particular example of murder is believed to be immoral, not only because of a shared belief that murder is wrong (*qua* disapproved of for various reasons) but because the particular act one is scrutinizing can be, and often is, interpreted more readily as an act of murder. Recall, that the same cannot always be said of alleged cases of poor taste. This is not necessarily because people vary with regard to what constitutes poor taste, although so doing would make a social norm more difficult to construct; rather, it is because, even if such a norm were to exist, the alleged example of trivializing murder may be ambiguous and therefore open to interpretation with regard to the realization of O. Consequently, while it may be much easier to achieve consensus in the case of "*this x* is immoral," especially where *x* is unambiguously identified as such (i.e., in the case of premeditated murder), consensus is less easily achieved in (many) cases of poor taste.

PUTTING IT ALL TOGETHER

If, as CEE attests, the process by which we arrive at attitudes congruent with accusations of immorality and poor taste is the same, then CEE needs to explain how we are able to differentiate, in a morally relevant way, between disapproval characteristic of immorality and disapproval characteristic of poor taste. In response, I have argued that, in the case of immorality, the standard against which something is deemed to be immoral is determined by the emergence of an intersubjective norm based on a shared *de re* attitude, even though different members of a society could have different reasons (*de dicto* attitudes) for holding this *de re* attitude. What each *de dicto* attitude has in common is that it is believed by the person who holds it to justify not only their negative *de re* attitude toward *x* but also the belief that the reason they have for holding this *de re* attitude (their *de dicto* attitude) is what makes *x immoral*. In other words, S_1, S_2,… S_n all believe that their respective reasons (*qua* their particular *de dicto* attitude) for having a negative *de re* attitude toward *x* is a *moral* reason for disapproving of *x* (stemming, say, from deontic or consequentialist principles, and so on) and is therefore what makes *x* immoral.[11]

The more these disparate societies agree on what their moral "requirements" are (to adopt a term used by Sinnott-Armstrong)[12]—as identified by their shared *de re* attitude toward them—the more universal the moral norm concerning what counts as a "requirement" becomes—although, again, the reason (*de dicto* attitude) for holding a particular *de re* attitude, and therefore identifying a particular "requirement," may vary both between and within these disparate societies. Like norms of morality, what constitutes poor taste within a society is established through consensus. But poor taste of the kind discussed, here, is *parasitic* on the morally pertinent, because establishing a norm with regard to the treatment of the morally pertinent (e.g., an established immoral activity) requires first and foremost a shared *de re* attitude toward that which is being treated in a particular way: that the object/event one is treating (inter alia) trivially/irreverently *is* (in this case) immoral.

Given that poor taste (as described) is parasitic on the morally pertinent, what constitutes a moral norm and, in virtue of this, what is morally relevant—whether moral or immoral—must be established first. Importantly, though, while this is a necessary condition for a norm of taste to be established, it is not sufficient. What is also required (a further necessary condition) is that consensus regarding the *interpretation* of that which is potentially in poor taste is achieved (i.e., people interpret the treatment of some morally pertinent matter in the same way). Thus, where one disapproves of O, given O's treatment of P (i.e., trivializing or misrepresenting P, and so on), there still needs to be sufficient agreement that *x* does in fact realize O (*qua* some token of O). Thus, where it is agreed that murder is immoral and even that trivializing murder would be in poor taste, it still needs to be agreed that *this* particular instance, even though it involves the representation of murder, is an example of someone or something *trivializing* murder. Recall from chapter 9, the winning entry at the Cornwall Beach Games sandcastle competition—a sand sculpture of a naked woman lying face down with a spade in her back, surrounded by police tape—and the different reactions it received.

Did the sand sculpture trivialize or in any other way make light of murder and/or the victims of murder? Should we interpret the actions of the police officers responsible for the sand sculpture as endorsing the view that actual murder is an appropriate subject for a light-hearted competition or for soliciting humor? According to the newspaper report, opinions differed on how to interpret the object and the police's action, and therein lies the problem. As noted earlier, it is much more straightforward to identify an immoral act like murder when it occurs (and agree that murder has occurred) than it is to agree how to interpret the *treatment* of a morally pertinent matter like murder, which has clear implications for any discussion on the morality of depictions of fictional immorality. Should agreement be forthcoming, however, then NI_5 (below) indicates the means by which one differentiates between poor taste and immorality:

> NI_5: *x* is not "something that ought not to be done because one disapproves (*qua* an expression of immorality) of P (e.g., violating God's law) and believes that *x* realizes P" but is "something that ought not to be done because one disapproves (*qua* an expression of offence, and therefore poor taste) of O, and believes that *x* realizes O in virtue of *x*'s perceived treatment of P (e.g., trivializing P)."

To further clarify the moral differentiation delineated within NI_5, we need to clarify the strength and nature of the "ought" used. In chapter 9, I set out the following difference between the "ought" intimated by one's moral pronouncement and that intimated by one's judgment about taste:

(a*) Pronouncing that "ϕ is immoral" intimates that one is required or obliged not to do ϕ.

(b*) Pronouncing that "ϕ is in poor taste" intimates that one "ought" (in some yet-to-be-determined sense, lying somewhere between recommendation and requirement) not to do ϕ.

I also argued that examples of poor taste are a species of suberogatory wrongdoing, and therefore that the sense of ought intimated by prohibitive pronouncement to do with taste is comparable to the strength of suberogatory prohibition: namely, something in-between recommendation and requirement. Ultimately, given that immorality and suberogatory actions are based on negative attitude, what separates them is the further belief concerning one's reason for the attitude: namely, whether it amounts to a reason one holds to be a moral reason or something else. The line demarcating immorality from poor taste or other suberogatory wrongs is fluid, as Durkheim's ([1895] 1982) example of a society of saints illustrates:

> Imagine a society of saints, a perfect cloister of exemplary individuals. Crimes, properly so called, will there be unknown, but faults which appear venial to the layman will create there the same scandal that the ordinary offense does in ordinary consciousness. If, then, this society has the power to judge and punish, it will define these acts as criminal and will treat them as such. For the same reason, the perfect and upright man judges his smaller failings with a severity that the majority reserve for acts more truly in the nature of an offense. (p. 68)

In sum, then, when examined fully, the proposition "*x* is not immoral but is in poor taste" amounts to the following noncontradictory conjunction:

> NI_6: *x* is not "something that must not be done" (on account of disapproving, *qua* an expression of immorality, of P and believing that *x* realizes P) but is "something that, in a suberogatory sense, ought not to be done" (because one disapproves, *qua* an expression of poor taste, of O, and believes that *x* realizes O in virtue of *x*'s perceived *treatment* of P, where P constitutes a morally pertinent matter).

CONCLUSION

In conclusion, what I hope to have shown, from the perspective of CEE, is that a morally relevant means of distinguishing between poor taste and immorality is available, thereby making the proposition "*x* is not in poor taste but is immoral" morally noncontradictory. Where both moral utterances and those concerning taste are attitudinal, and where the attribution of one's attitude involves the same underlying process, a moral contradiction is avoided because taste expresses one's attitude toward the treatment of something one (one's society) already considers to be immoral.

Validating the distinction between poor taste and immorality affords a more nuanced approach to the morality of fictional immorality. In the final chapter, I utilize this distinction to delineate the wrongness of depictions of fictional immorality and, in doing so, summarize the conditions under which fictional immorality, and/or our engagements with it, ought to be classified as immoral or something approximating a suberogatory wrong.

NOTES

1. Reprinted by permission from Springer Nature: Springer Nature, *Ethical Theory and Moral Practice*, (Young, Garry. An Expressivist Account of the Difference between Poor Taste and Immorality. *Ethic Theory Moral Prac* 22, 465–482 [2019]. https://doi.org/10.1007/s10677-019-09998-2), Copyright © 2019, Springer Nature (2019).

2. See Ridge (2006) for discussion on the forerunner to CEE—namely, *ecumenical expressivism*—to which CEE is indebted.

3. By way of an example, a *World Wildlife Fund* advertisement was condemned for being offensive and tasteless because of its alleged downplaying of the 9/11 attacks; see https://www.telegraph.co.uk/news/worldnews/northamerica/usa/6131150/WWF-advert-condemned-for-downplaying-911-attacks-on-New-York.html. Similarly, in Australia, the "Operation Bomerang" advertisement—designed to coincide with the Australia Day celebrations—was considered to be in poor taste by some. https://www.smh.com.au/entertainment/operation-boomerang-antivegan-australia-day-lamb-ad-featuring-lee-lin-chin-gets-the-allclear-20160120-gm9yqm.html (all accessed November 4, 2019).

4. By way of an illustration of the latter example, consider the following joke, which some *may*(?) find to be in poor taste. A doctor, a lawyer, and a fundraiser arrive at the pearly gates of heaven. St. Peter tells the doctor that he will grant him one wish before he enters heaven, so the doctor asks for a million dollars. St. Peter grants the wish and the doctor enters heaven. This generosity did not go unnoticed by the lawyer so when St. Peter asks him for his wish, the lawyer asks for a billion dollars. St. Peter grants the wish and the lawyer enters heaven. When St. Peter asked the fundraiser what she would like, she says, "If it is not too much trouble, could I please get the business cards of the two people who entered heaven just ahead of me?" Taken from: http://www.bemonsterful.com/index.php/easyblog/entry/fundraising-humor (accessed October 12, 2019). Of course, one might argue that the joke is not intended to target fundraising *per se*, but the perceived "pushy"/aggressive nature of some fundraisers.

5. To further clarify, the term a "kind of objectivity" means simply that the moral standard is externalized beyond our individual (subjective) preferences (Goodwin & Darley, 2012; Stanford, 2018) and is therefore independent of any one individual; yet, the moral standard is not completely mind-independent (Brey, 2003; Searle, 1995) in the sense required for moral objectivism.

6. It may be that S has a negative attitude to murder because she believes it violates one's duty to others *and* creates more harm than good (negative utility) *and* is a vice: all things of which S disapproves. It is likely that S will prioritize one of these over the others, although not doing so does not undermine CEE.

7. This issue is part of a larger challenge for expressivist approaches to morality: namely, the *moral attitude problem* (Miller, 2003), whereby it is claimed that an expressivist approach lacks the resources to differentiate between a negative, nonmoral attitude—say, one's dislike of one's favorite team's new away strip—and a negative moral attitude toward cheating in sport. In the latter case, it is (allegedly) unclear which specific kind of conative attitude constitutes moral thinking, specifically (which is why it is also known as the *specification problem*; see Björnsson & McPherson, 2014). It is, however, beyond the scope of this book to address the moral attitude problem in its more fundamental form (for a detailed discussion, see Köhler, 2013, who argues that the "problem" is not unique to expressivism). Instead, I seek to tackle a particular subspecies of this problem: namely, differentiating between attitudes constitutive of poor taste and immorality, respectively.

8. Smuts (2016) likewise distinguishes between robust responsibility in the case of an immoral act like murder and disesteem as a reaction to a lesser moral infraction.

9. Given the position I am adopting, it could be that S_3 disapproves of murder and considers it to be immoral because they *believe* that murder violates the instruction of the alien prophet Zog. As already noted, I do not consider the diversity of reasons (*de dicto* attitudes) to be a weakness of CEE. The fact that different reasons can be given for why something is immoral (based on different beliefs about what counts as a measure of morality) means that the shared belief that *x* is immoral is more robust, given that a number of reasons for why it is immoral would have to be challenged in order to challenge the overall claim, even where the validity of some reasons may be more easily disavowed than others.

10. Saying this does not rule out the possibility that one could find it offensive despite being amused by it or that one could be amused by it while recognizing that it would likely offend others (see Woodcock, 2015, for a detailed discussion on these and related points).

11. For a detailed discussion on different forms or components of moral reasoning, see Saunders (2015).

12. Sinnott-Armstrong (1987, p. 265) refers to those things "it would be morally wrong not to act on without any moral justification or excuse" as *requirements*.

Chapter 14

Immoral Fiction and Censorship

There is no such thing as a moral or an immoral book. Books are well written, or badly written. That is all.

—Oscar Wilde, preface to *The Picture of Dorian Gray*, 1891

INTRODUCTION

The implication of Wilde's famous assertion that books should be subject to aesthetic and not moral scrutiny[1] is at odds with the aim of this book, which is to uncover the conditions under which fictional immorality counts as immoral fiction. In accordance with this aim, when addressing the question "What are the necessary and sufficient conditions for the moral condemnation of depictions of fictional immorality and our interactions with them?" a number of "contender conditions" were examined, such as content, meaning, and (potential for) harm, and one's unwillingness to fictively imagine the fiction, the manner of one's enjoyment, and whether it features depictions of pure fictional or actual (including historical) figures and events. I also considered whether the fact that one's fictive imagining was in the form of (private) fantasy mitigated any charge of immorality. Moreover, in an attempt to present a more nuanced account of moral wrongdoing—within the expressionist approach I favor (namely, CEE)—I differentiated between what it means to say that one ought not to do *x* when it is in poor taste (*qua* a suberogatory wrong) compared to when it is immoral. I will have more to say about this distinction in the section "From Moral Condemnation to Censorship" when discussing immorality and moral censorship.

The aim of this final chapter is twofold: (i) to summarize the conditions under which fictional-x_i and our engagements with it are immoral or, if not immoral, constitutive of a suberogatory wrong, and (ii) to discuss the issue of moral censorship and its relation to fictional content judged to be immoral or in poor taste.

TAKING STOCK

What can we conclude about the conditions under which fictional immorality becomes immoral fiction? Starting with content, I rejected the assertion that the mere fact that the fiction depicts that which is immoral is sufficient for it to be immoral. In short, content depicting an immoral act that is fictional is not, in and of itself, sufficient for a claim to immorality. Instead, I argued that, *tout court*, there is no logical or conceptual entailment between the immorality of x and the immorality of fictional-x_i. Therefore, I reject not only the universality of M1 but also M3's selective condemnation, even in cases with strong(er) intuitive appeal, such as virtual (or otherwise fictional) pedophilia and various discriminatory slurs. Support for M2—the judgment that there is no relationship between the immorality of x and the immorality of the content of fictional-x_i—does not negate, of course, the possibility that other sufficient conditions exist for the ascription of immorality besides fictional-x_i's content, such as whether the fiction endorses an immoral worldview or is ambiguous in this regard or whether, irrespective of meaning, engaging with the fiction causes (or increases the likelihood of causing) harm.

As far as the meaning of fiction is concerned, I support IWV: that a depiction of fictional-x_i is immoral *if* it is congruent with the fiction's overall endorsement of an immoral worldview. I also acknowledge that it is one thing to support IWV and quite another to apply it to a specific example of fictional-x_i, given that it is unlikely that a fiction that seeks to endorse the immorality of what it depicts will always do so explicitly and unambiguously. In recognition of this, I entertained the possibility of "erring on the side of caution" and adopting a revised version of IVW based on a *reasonable interpretation* of the fiction's meaning. I rejected this revision, however: holding that the ontological claim that fictional-x_i *is* immoral cannot be deduced from its epistemic premise, no matter how reasonable the interpretation on which the conclusion is based. Justification for the claim that fictional-x_i is immoral therefore requires a greater weight of evidence than can be amassed by a *reasonable* interpretation, which I accept may not be easy to procure. I did, however, propose that the "reasonable interpretation" approach, and therefore erring on the side of caution, could justify the claim that fictional-x_i is morally insensitive rather than immoral, particularly where the depiction aligns with

Patridge's (2011) notion of incorrigible social meaning or other sensitive topics. The video game *Jesus Strikes Back*, for example, depicts various famous/infamous individuals—inter alia, Jesus Christ, Donald Trump, Vladimir Putin, and the Christchurch mosque gunman—engaged in violent acts, often against minorities.[2] In the absence of definitive evidence of the endorsement of the immorality depicted within the gameplay, taken as a whole, it nevertheless seems reasonable to hold that what is depicted—at least in the case of the Christchurch gunman or Jesus, dressed as a crusader, killing Muslims—is somewhat insensitive and akin to poor taste. That said, I countered this claim by arguing that the intention to morally educate or cultivate a greater aesthetic appreciation could mitigate charges of moral insensitivity and poor taste. (As an aside, I am not suggesting that *Jesus Strikes Back* possesses the aforementioned mitigating qualities.)

As for the occurrence of harm, either as a direct or as an indirect consequence of fictively imagining depictions of fictional-x_i, and either toward oneself or toward another, while I support, in principle, the idea that engaging with fictional-x_i is immoral *if* it causes harm (non-*pro tanto* harm, that is), empirical support for a direct or indirect causal connection between fictively imagining depictions of fictional immorality and harm is mixed, resulting in an absence of consensus over the effects of engaging with fictional-x_i. Given this, the jury is still out on the question of harm.

I next proffered a precautionary approach, whereby engaging with fictional-x_i is immoral *if* it *risks* (*qua* increases the likelihood of) causing harm, but concluded that this was problematic. First, because it is not immoral to engage in a risk-increasing activity if one consents to this. Second, because even if one is of the opinion that, in the absence of consent, it is immoral to engage in an activity that increases the risk of harm to others, the application of this moral outlook to depictions of fictional immorality and one's engagement with them suffers from the same problem of contradictory and contested empirical evidence that plagues the causal account: there is no consensus that engaging with fictional-x_i is a risk-increasing activity, regardless of medium (e.g., film, TV, video game).

Following this, I argued that the immorality of enjoying fictional-x_i is dependent on the type of enjoyment one seeks and, also, as a consequence, the type of game world one creates in order to elicit this enjoyment. Where one seeks to enjoy the depiction itself (for its own sake), then such enjoyment (*qua* enjoyment$_{(sim)}$) is not sufficient for a charge of immorality. If, instead, one uses the depiction as a vehicle to satisfy vicariously what the fiction is a depiction of (e.g., actual murder or rape or torture), then the enjoyment elicited (*qua* enjoyment$_{(sub)}$) is immoral, as is the game world created to elicit this enjoyment$_{(sub)}$. Where one is unwilling to fictively imagine a particular game world, however—say, in the case of a fiction depicting the make-believe morality of racial or sexual discrimination—then this unwillingness is not sufficient evidence for immorality (whether in terms of content or engagement). The relationship between the belief that *x* is immoral and the belief that a depiction of fictional-*x* is immoral, such that one is unwilling to even imagine it, lacks the kind of coherence discussed in section "The Compatibility of the Moral Claims Made about Fictional-x_i with CEE." Imaginative resistance therefore struggles to establish any normative credentials. That said, where one's unwillingness to imagine fictional-x_i stems from, say, the belief that the fiction is endorsing an immoral worldview or risks causing harm, then, while this could in principle justify a claim to immorality, it is still subject to the same problems of empirical verification already discussed in relation to harm and endorsement. Because of this, as a psychological phenomenon, imaginative resistance may be a more reliable indicator of depictions that are morally insensitive or, perhaps more assuredly, be a sufficient means of indicating personal tolerances, at least as far as depictions of fictional immorality are concerned.

Thus far, we can conclude that fictional-x_i is immoral *if*:

1. It depicts immorality congruent with an immoral worldview endorsed by the fiction taken as a whole.
2. Engaging with it (*qua* one's fictive imagination) causes (a non-*pro tanto*) harm to oneself or another.

In addition:

3. *If* one enjoys$_{\text{(sub)}}$ fictional-x_i, then one's enjoyment is immoral.

Interpreting conditions 1–3 through the lens of CEE means that the immorality of *x* (e.g., establishing whether murder or racial discrimination is wrong) is determined by the construction of an intersubjective norm. Given the existence of this norm, in relation to condition 1, as discussed previously, it follows that *endorsing* the worldview that *x* is moral, or at least not immoral (even through the medium of fiction), is itself immoral because it runs contrary to the constructed norm concerning which worldviews are moral. Depictions congruent with that worldview are therefore immoral *if* (and therefore because) they amount to the material by which the worldview is being endorsed. To be clear, this is not because of their content per se but because of their meaning. Moving on to condition 2: where the intersubjective norm exists that causing (non-*pro tanto*) harm is wrong, it follows that a depiction of fictional immorality that causes this type of harm is immoral. As for condition 3: where the intersubjective norm establishes what is immoral, then this particular way of enjoying fictional-x_i—enjoying$_{\text{(sub)}}$—is in effect endorsing immorality and so, in accordance with condition 1, is immoral.

Conditions 1–3 can be supported, in principle, irrespective of the ease with which they are implementable. Moreover, these sufficient conditions apply to both pure fictional immorality and depictions that feature actual or historical figures and/or events. In the latter case, I have also argued for additional moral concerns beyond 1–3, such as harm to one's reputation and a lack of respect, both of which are particularly pertinent to a relatively new form of fiction known as deepfake pornography. (As an aside, the immorality of nonconsensual pornography, discussed in relation to deepfake pornography, is also something that is more applicable to this new form of "fiction" rather than other fiction discussed.) Harm to one's reputation can be subsumed under the more general principle regarding harm (at least where the implied understanding is that such reputational harm is undeserved and therefore illegitimate).[3] Therefore, in addition to conditions 1–3, in the case of fictional immorality featuring actual (living) persons:

4. Fictional-x_i is immoral *if* it is created for the purpose of disrespecting an actual person's inherent value as a person (as opposed to, say, their status or social standing), irrespective of its success and any subsequent harm caused.

Again, interpreted through the lens of CEE, where the act of disrespecting someone—*qua* not valuing them as a person—exists as a constructed moral norm, and where fictional-x_i is believed to realize a property disapproved of by those constitutive of the intersubjective norm (in this case, in virtue of the belief that it disrespects the person targeted), then fictional-x_i, at least in terms of the use for which it was created or is now being put, is immoral.

As noted, condition 4's application to deepfake pornography is particularly apt but need not be restricted to this. Imagine that the screenplay to *Amadeus* had been penned during Salieri's lifetime. Would such a play satisfy condition 4? Possibly. But to do so, it would need to be established that the fiction was created (or was being used) intentionally to disrespect Salieri. Recall the Buscemi-Lawrence (nonpornographic) deepfake: while this fiction may perturb and/or even offend some, it is debatable whether it would satisfy condition 4.

Conditions 1–3 apply to surrogate fantasy in the following ways: Condition 3 is directly applicable. As for condition 1, even though the fantasy may be limited in scope, and may indeed amount to the fiction taken as a whole, it is nevertheless congruent with the fantasist's worldview that, at the very least, the object of fantasy is someone with whom the fantasist would like to act immorally outside of the fantasy (e.g., sexually assault), even if the desire is never acted on. The fantasist's worldview is therefore immoral to the extent that, minimally, it includes the belief that it acceptable to desire to sexually assault someone, or perhaps specifically *this* person. By willingly creating and/or embellishing the fantasy, the fantasist endorses their own immoral worldview: a view that must have existed prior to the *surrogate* fantasy. For in order for this type of fantasy to have been *willingly* created and/or embellished, the worldview it endorses must have existed, or been disposed to exist, in the first place.

In relation to condition 2, where the surrogate fantasy is an expression of one's preexisting immoral beliefs and desires, the extent to which the fantasy itself does any more harm than is already occurring in virtue of one's preexisting immoral beliefs/desires is not something that can be established a priori. In other words, whether the fantasy helps cultivate the preexisting beliefs and desires, and/or increases the risk of carrying out the fantasy for real, is an empirical matter. What can be established, a priori, however, is that the existence of beliefs and desires that fuel the fantasy harms the object of one's fantasy because X (*qua* object of the fantasy) is treated as someone toward whom one holds immoral beliefs and desires.

Condition 2 is likewise applicable to idle fantasy: for, in the case of idle fantasy, as with all fictive imagining, the extent to which engagement causes harm to oneself or others, to the point of allowing an inference to be made from the particular to the general, is an empirical matter awaiting sufficient research support. To be clear, though, in the case of idle fantasy, there are no preexisting immoral beliefs of the kind that grounds certain surrogate fantasies to fuel this particular fantasy. The possible existence of the belief that X is a suitable object of *idle fantasy* is, however, problematic, insofar as a case can be made for a lack of respect; although, as I have argued, the lack of a public expression of disrespect may mitigate the wrongdoing to some degree. Certainly, in the case of condition 4, there is perhaps an assumption that the disrespect intended is synonymous with public expression; and while this is not a requirement for the wrongdoing of disrespect, as the case of idle fantasy demonstrates, it does muddy the moral waters somewhat regarding the severity of wrongdoing. It is possible, then, that the disrespect constitutive of the sorts of idle fantasies discussed amounts to an act of disesteem and is therefore tantamount to a suberogatory wrong.

The inclusion of suberogatory wrongs allows the occurrence of moral wrongdoings that we are not obliged to desist from performing—meaning they are not immoral. We saw an example of this when discussing moral insensitivity, particularly in relation to poor taste (i.e., when a reasonable interpretation of the fiction could not determine whether the content was depicting or endorsing an immoral worldview). In the absence of mitigation, such as moral edification and/or aesthetic appreciation, fictional-x_i can be classified as a suberogatory wrong in virtue of its moral insensitivity; or, if the depiction concerns the offensive treatment of a morally pertinent matter, poor taste. As was discussed with the John Waters example, aesthetic quality could mitigate the lack of moral taste, transforming what Waters calls *bad*, bad taste into *good*, bad taste. A charge of moral insensitivity and/or poor taste could also be leveled at someone for enjoying$_{\text{(sim)}}$ fictional immorality, particularly if it depicts something that has incorrigible social meaning.

The conditions identified for the immorality of fictional immorality are all sufficient conditions. Are any conditions necessary? In a trivial sense, yes. It is trivially necessary that ϕ, *qua* immoral fiction, satisfies the necessary (and sufficient) conditions for ϕ to be a fiction (as presented in chapter 2) and, in the case of depictions of fictional immorality, necessarily depicts fictional immorality. Nontrivially, perhaps, the closest to a necessary condition is a fictional depiction of immorality that causes harm. Such an all-encompassing notion of "harm" is, however, somewhat nebulous. It would not be necessary, for example, for the harm to be physical; neither would it be necessary for it to be emotionally or in some other way psychologically detrimental. The direct causal path of the harm, and, subsequently, its proximity to the one engaged with the fiction, is also unclear (recall the discussion on harm to self and others in chapter 5). Moreover, if one is prepared to endorse the idea of a harmless wrongdoing (see chapter 6), whereby something can be materially innocent (not directly or proximally harmful) but morally non-innocent (nevertheless wrong in some sense), as well as the possibility of a more protracted and distal cultural harm, then the concept of harm itself becomes too broad to be informative as a necessary condition. Consequently, it is my contention that there are no nontrivial necessary conditions for the immorality of fictional immorality.

Having identified sufficient conditions for the immorality of fictional-x_i, or if not immorality, then a suberogatory wrong, what, if anything, justifies the moral prohibition—*qua* censorship—of morally problematic fiction?

FROM MORAL CONDEMNATION TO CENSORSHIP

Bourke (1956) proposed that censoring is "the act of restricting the public expression of thoughts, feelings or volitions . . . considered by the censoring agency to be contrary to public or private good" (p. 58). This definition will suffice for our discussion (for more detailed examination of definitions of censorship, however, see Moore, 2013; O'Leary, 2016).

There are a number of ways in which censorship can be applied and a diverse set of reasons to censor: inter alia, classified information censored for reasons of national security or militarily sensitive information censored from certain communications during wartime (Irving & Townend, 2016; Robins & Webster, 1986).[4] These forms of censorship are not of concern, here, however. Instead, I am interested in censorship as an act of moral prohibition, and subsequently whether the fact that something is immoral or in some other way morally problematic—such that, in some sense, it ought not to be done—is sufficient to justify censorship; or whether, and without fear of contradiction, one can adopt the view that fictional-x_i is immoral or a suberogatory wrong but reject the idea that it should be censored because of this.

Censorship and Harm

Attempts to justify censorship often present as follows: (i) as protection from harm and (ii) as a way to preserve freedom (McCormick, 1977). Arguments for the justification of censorship in art and fiction more generally, based on protection from harm, tend to overlap with legal justifications for the prohibition of obscene material which, explicitly or implicitly, likewise draw on arguments from harm (recall discussion in chapter 6 on moral corruption and appeals to prurient interest). Where it can be shown that fictional-x_i causes harm to oneself and/or others—that is, where a clear causal connection has been established, from the particular to the general, between engagement with fictional-x_i and a direct form of physical and/or emotional and/or other psychological harm—then an argument in favor of the censorship of this material, for this reason, should prove persuasive if not irresistible.

As already discussed, however, empirical findings on the effects of engaging with fictional-x_i are inconsistent, and there is a lack of consensus regarding what the research evidence shows. Moreover, the notion of harm, as mentioned earlier, can be nebulous, and in some cases, causing harm may even produce a good outcome. To illustrate: Suppose, I make a political statement against an oppressive regime through the use of fictional immorality that harms the regime by weakening it, politically (McCormick, 1977). In such a situation, the harm caused to the regime is likely to be perceived (by some/many?) as a morally good thing. As things stand, then, unless it can be convincingly shown that censoring x negates, or would negate, an intrinsically bad harm, the more general position that such action is intended to protect us from suspected harm, or the *risk* of harm, is insufficient to justify censorship.

Suppose, instead, I judge a poem to be immoral because it includes fictional content that is racially charged and I am told, on good authority, that what is depicted reflects the racist views of the poet.[5] If true, then it is immoral because it satisfies IWV. But even if it is immoral for this reason, which does not require a direct causal connection to harm—although one may be of the opinion that the meaning underlying the poem could increase the likelihood of indirect harm occurring by inciting people to behave in discriminatory ways (see below for a return to this point)—should this be reason enough to censor it? I do not believe so, and I do not believe so even though I have argued that pronouncing that ϕ is immoral intimates that ϕ ought not to be done—meaning that we are morally obliged not to do it.

I see no contradiction between these two positions. Declaring that ϕ ought not to be done is to prescribe a way of acting (which, of course, implies a way not to act). Should one choose to violate this moral obligation, then what one has chosen to do is immoral (in accordance with certain moral norms) and needs to be identified as such and condemned. Strictly speaking, censorship of fiction, as an expression of moral prohibition, does not prevent one from depicting fictional immorality deemed immoral (i.e., it does not make one incapable of doing this); rather, it seeks to restrict others' access to the material. In other words, it seeks to limit its public expression (Bourke, 1956); and it does so, if not to protect from (alleged) harm, then to preserve our freedom from certain forms of harm, such as persecution or deceitful misrepresentation. But in order to do this, censorship necessarily denies a freedom of its own: namely, the freedom to express oneself as one sees fit in more publicly accessible ways.

Censorship and Freedom of Expression

Bourke (1956) holds that we have certain inalienable rights, such as the right to freedom of expression. Given this, censorship would appear to oppose this right, at least when *exercising* it is deemed to have negative consequences (i.e., when what is expressed runs *contrary to the public or private good*, to borrow Bourke's phrase from earlier). If censorship is to be considered legitimate, then one's right to freedom of expression seems to be contingent on the felicitousness of consequence (Ward, 1990). But by distinguishing between the inalienable right itself and one's freedom to *exercise* that right, the tension between freedom of expression and censorship is eased. Thus, while I possess the inalienable right to express my thoughts, feelings, and behavior, the *exercise* of this right—say, through the enactment of fictional immorality—may at times be constrained, owing to the harm it causes or, as some might argue, is likely to cause, at least where the harm is intrinsically bad. To draw on a well-used example: Refusing to allow someone to yell, "Fire!" as a joke in a crowded theater is not a violation of their *right* to free expression, only to its indiscriminate exercise. Thought about in this way, it is not inconsistent to hold that "people retain the right to free expression even when a government is justified in preventing them from exercising this right" (McCormick, 1977, p. 37).

In the next section, I discuss some of the points raised above, and evoke CEE, rather than rights, as a defense against moral censorship.

CEE AND CENSORSHIP

CEE is a form of moral relativism that allows—as in, makes it legitimate for—one *de re* moral attitude to be prioritized over another, based on robustness of attitude (i.e., the greater the number of *de dicto* attitudes that contribute to the formation of the intersubjective *de re* attitude, the more robust the moral norm); and while CEE promotes the view that the more robustly constructed the attitude, the less likely it is to change, it nevertheless recognizes that change is possible. Morality, under CEE, is conceived as fluid rather than absolute, but the robustness of its construction makes changes to moral norms less capricious. Censorship hinders CEE's notion of robust relativism by reducing the diversity of attitudes publicly available against which the intersubjective norm is to be tested and, ultimately, from which it is derived.

Similar support for a cauldron of ideas, against which one tests (as a means of justifying) a belief, is presented by Mill (2005 [1859]) in his argument against censorship. To paraphrase Mill:

1. No governing body is infallible. Given this, it is better to test a contrary view than censor it. In this way (as Shakespeare informs us), "truth will out."[6]
2. Even if a contrary position is false, there may be an element of "truth" contained within the erroneous argument used to assert or defend the contrary position that, in accordance with 1, needs to be uncovered. (By "truth," I mean simply that which would contribute to our understanding and thereby increase an established position's verisimilitude.)
3. A position that is held without being tested is a position that lacks justification. To avoid the charge that one is embracing dogma, it is not enough that one believes a given position, one must be willing and able to justify one's belief against contrary views, rather than insulate it from them.
4. In the absence of debate and the rigorous rebuttal of challenges, any doctrine, whose virtues one seeks to extol, risks losing its meaning. If a position is reduced to dogma, then one's *understanding* of it—its merits and therefore why it should be upheld—will be eroded. In short, through censorship, and therefore in the absence of the need for a defense of one's belief, there is a danger that the meaning of established doctrine (including one's understanding of it) will be reduced to "it just is."

In line with points 1 and 2, recall that it is not a requirement of CEE that the property S *believes x* realizes (about which S disapproves) is actually a property *x* realizes. In other worlds, CEE does not require that what S believes about *x* is true. A recognition of infallibility is therefore built into CEE. Instead, CEE requires only that, whatever belief S holds about *x*, it is something she is capable of testing and, if required, revising irrespective of whether she is willing to do so.

Dworkin's (2006) counters Mill's appeal to tolerance (i.e., that we should tolerate the existence of alternate views because debating them brings us closer to the truth) with a rhetorical question: "How can we expect people who are committed to a particular faith, as a value transcending all others, to tolerate its open desecration?" (Dworkin, 2006, p. 130). Whether they are *willing* to tolerate a challenge to their worldview, especially one perceived to be a form of desecration, is an empirical matter perhaps best left to psychologists and other social scientists to determine and explain. Under CEE, moral agents, as rational agents, are assumed to be *capable* of doing this, at least for the purpose of critical examination; but, more than this, they are required (irrespective of willingness) to exhibit tolerance for the *sake* of critical examination which itself is instrumental to a *robust* outcome.

For Mill, they ought to do this as a means of avoiding dogma: by strengthening their justification for what they believe. Likewise, for CEE, the existence of divergent attitudes should be tolerated in order to have the best "materials" available from which to construct a more robust *de re* attitude and hence moral norm. To be clear, tolerance of this kind is not equivalent to ascribing such views equal justification and moral status; one could still declare that *x* is immoral and be justified in doing so. Rather, such tolerance is grounded on the idea that knowledge of other beliefs strengthens the case for one's own beliefs *if* they can fend off the challenge posed by alternatives and, in the case of CEE, strengthen the intersubjective norm that is constructed *and justified* as a result of testing these different views in order to determine which are the most coherent and empirically verifiable.

But what of the argument that censorship in fact protects truth? To argue this is to disregard centuries of evidence to the contrary, particularly where the "truth" one is safeguarding is contentious (e.g., a religious view). Moreover, removing from consideration an alternate belief is a perverse form of protection. Certainly, it does not strengthen one's own case (although it may safeguard it); rather, it risks stunting its growth toward verisimilitude for the reasons outlined by Mill, which, I maintain, are compatible with the aim of CEE's robust relativism (i.e., verisimilitude for CEE is increasing the *robustness* of one's socially constructed and therefore objectified norm—a position that cannot be arrived at by default in virtue of reducing the options available for examination).

As for the censorship of make-believe, *qua* fiction and other art forms, for Sierz (2001), it negates or at best hinders the artist's opportunities to take risks—say, by exploring alternative views and expressions of morality and/or criticizing conventional wisdom—by depriving such alternative views of a more publicly accessible outlet for their expression.

Mill presents a consequentialist defense of freedom of expression by arguing that censorship has a negative effect on our relationship with truth (Funk, 1984). By its very nature, censorship hinders, or at least does not facilitate, the discovery of truth and so one ought not to permit (in a moral sense) that which produces a net negative consequence. In summarizing Mill's position, Ward (1990) states: "Censorship is wrong because it makes it less likely that truth will be discovered . . ., and it is wrong because it has a destructive consequence for the intellectual character of those who live under it" (p. 86). Ward (1990) also notes, this time from a deontological perspective, that censorship should be judged impermissible because it restricts the flow of information and ideas: something that is fundamental to our nature as rational and autonomous human beings.

Both the consequentialism of Mill and the deontological argument characteristic of Kant's categorical imperative present different justifications for what Bourke (1956) refers to as *freedom from* something. In this case, freedom from the constraint of censorship. But Bourke also holds that this type of freedom is secondary to a more fundamental freedom: the *freedom for* something. Thus, to be free to express oneself requires not only the absence of restraint, *qua* freedom from censorship (in this case), but also the freedom to develop the skills to express oneself effectively courtesy of, say, access to education, or the time for creative and/or leisurely pursuits and/or social engagements and free association.[7]

CEE necessarily favors freedom of expression because its capacity to present a *robust* form of moral relativism depends on the existence of publicly expressible alternate views. It therefore requires freedom for those conditions that facilitate free expression (e.g., access to education, free association, and so on). But, as previously noted, free expression also requires freedom from censorship. More formally:

- Robust relativism *only if* freedom *from* censorship, freedom *for* the conditions that facilitate free expression, and free expression

If CEE, as an advocate of robust relativism, requires (*qua* a necessary condition) freedom from censorship, then how are we to understand CEE's claims that fictional-x_i is immoral? After all, such a claim intimates that fictional-x_i ought not to be done, which suggests censorship. In other words, if fictional-x_i ought not to be done (because it is immoral), then why not censor it? For CEE, the act of censorship is a double-edged sword. On the one hand, censorship provides a means of acting against depictions of fictional immorality that are deemed immoral, thereby curtailing accusations that CEE is impotent. On the other hand, censorship appears to diminish the confidence we can have in CEE to make robust moral decisions. As a consequence, for

CEE, it is ultimately self-defeating. As champion of CEE, I am therefore of the opinion that freedom of expression is inalienable. To say this, however, is not to advocate the indiscriminate *exercise* of this freedom. Declaring that fictional-x_i is immoral—say, because it accords with IWV—is therefore not sufficient to negate the exercise of free expression through an act of censorship; nor is the immorality of enjoying$_{\text{(sub)}}$ fictional-x_i sufficient to restrict S's access to fictional-x_i. Instead, where fictional-x_i and/or one's engagement with it is deemed immoral:

> C_1: Censorship of fictional-x_i *iff* the absence of censorship prevents the realization of at least one of the conditions required for free expression

If one's freedom to express racially charged fiction inflicts on the minority group conditions that directly bring about a reduction in *their* freedom of expression (say, by causing a reduction in their education and skills training, or a reduction in the time/opportunity afforded them to be creative or to associate freely), then the racially charged fiction should be censored, in accordance with C_1. What is not clear, of course, is how a racially charged fiction in and of itself would have the potency to cause such social upheaval. Perhaps its power lies in the creative way it captures a preexisting racist point of view: perhaps by making the view more appealing to a wider audience, superficially at least, by reinforcing negative stereotypes. Such a scenario is more plausible than some seemingly mystical transformation of social norms (into a kind of apartheid system) through the power of the fiction alone (because such fiction would normally depend on preexisting immoral attitudes for purchase). Even so, where a fiction is used with the intention of encouraging social change through (for example) the promotion of discrimination along racial lines (i.e., by reinforcing already-existing racist attitudes and/or behavior), such is the nature of CEE that even a view that is deemed abhorrent by current standards must be permitted in order that our current moral norms can be pitted against it. Not only to undermine any threat posed by the immoral view endorsed by the fiction but to stave off the charge that current norms have been reduced to, or are fast becoming, dogma.

Censorship should be rejected, then, *unless* it can be shown that the immoral fiction has had, or will have, a direct and immediate causal effect of the kind described above, which must move beyond mere moral outrage. This is unlikely to occur, however, owing to the fact that the fiction will most probably be endorsing already accessible racist views, which must be taken into account when determining the causal efficacy of the fiction, alone, to produce the conditions required to satisfy C1.

In sum, under CEE, declaring that fictional-x_i is immoral is not sufficient to justify censoring the fiction, even when a claim to immorality intimates that fictional-x_i ought not to be done and one's moral condemnation is intended to notify its target of this fact. It is not sufficient to justify censorship because censorship would hinder CEE's ability to construct robust moral norms and thus become self-defeating. The only exception to this position is where a lack of censorship prevents at least one of the conditions necessary for freedom of expression to be realized. Censorship of this kind would not affect CEE's ability to construct robust moral norms (*de re* moral attitudes) because what is being censored is a condition that itself would prevent the construction of robust moral norms.

CONCLUSION

What I hope to have identified and defended throughout this monograph are the sufficient conditions for the immorality of moral fiction (any necessary condition identified being regarded as only trivially so). It is further hoped that the sufficient conditions presented in this final chapter, and defended throughout the previous chapters, can be used as a moral framework to inform future debate on the morality of depictions of fictional immorality and perhaps even legislation, including censorship of the kind exercised by the likes of the Australian Classification Review Board (ACRB) and the British Board of Film Classification (BBFC).[8]

I would like to finish by acknowledging that any arguments grounded on current research findings are contingent on these findings and their interpretation by leading scholars in the field and therefore that the conclusions drawn about the relationship between, for example, engaging with media violence and (the likelihood of) harm are subject to further research outcomes.

NOTES

1. Although Wilde's comment could arguably apply to all books, he probably had novels (and therefore fiction) in mind when he wrote it. It was, after all, written in the preface of the unabridged version of his fiction *The Picture of Dorian Gray*, as part of his response to those who accused the initial, shorter version of his novel—which first appeared in 1890 in Lippincott's monthly magazine—of being immoral.

2. https://www.newsweek.com/anger-over-sick-video-game-that-allows-you-play-trump-gunning-down-migrants-feminists-antifa-1441745 (accessed January 22, 2020).

3. I am not suggesting that, where harming a reputation is deserved, a pornographic deepfake constitutes a legitimate means of doing so. After all, such a deepfake would still be a form of nonconsensual pornography, and therefore immoral because of it.

4. Shusterman (1984) proposes “aesthetic censorship” whereby *x* is censored on aesthetic grounds: namely, for being deleterious to the goal of achieving a higher aesthetic standard. Of course, one might also argue that a moral flaw can affect the aesthetic standard of a work of art or other fiction (see discussion on moralism/ethicism in Gaut, 2007).

5. The poetry example is inspired by the following real-life example: https://www.bbc.com/news/world-europe-45717680 (accessed January 23, 2020).

6. *Merchant of Venice*, Act II, Scene 2.

7. Ward (1990) draws a similar distinction when discussing negative and positive rights.

8. When discussing censorship, I do not seek to reject the need for age restrictions and therefore age-related censorship.

References

Abraham, W. E. (1974). Disentangling the “cogito.” *Mind*, *83*(329), 75–94.

Adams, A. A. (2010). Virtual sex with child avatars. In C. Wankel & S. Malleck (Eds.), *Emerging issues in virtual worlds* (pp. 55–72). Charlotte, NC: Information Age Publishing.

Al-Alosi, H. (2018). *The criminalisation of fantasy material: Law and sexually explicit representations of fictional children*. London: Routledge.

Ali, R. (2015). A new solution to the gamer’s dilemma. *Ethics and Information Technology*, *17*(4), 267–274.

Allen, K. (2015). When is a slur not a slur? The use of nigger in “pulp fiction.” *Language Sciences*, *52*, 187–199.

Alward, P. (2010). That’s the fictional truth, Ruth. *Acta Analytica*, *25*, 347–363.

Anderson, C. A. (2004). An update on the effects of playing violent video games. *Journal of Adolescence*, *27*, 113–122.

Anderson, C. A., Berkowitz, L., Donnerstein, E., Huesmann, L. R., Johnson, J. D., Linz, D., & Wartella, E. (2003). The influence of media violence on youth. *Psychological Science in the Public Interest*, *4*(3), 81–110.

Anderson, C. A., Bushman, B. J., Bartholow, B. D., Cantor, J., Christakis, D., Coyne, S. M., et al. (2017). Screen violence and youth behaviour. *Pediatrics*, *140*(2), 142–147.

Anderson, C. A., Carnagey, N., & Eubanks, J. (2003). Exposure to violent media: The effects of songs with violent lyrics on aggressive thoughts and feelings. *Journal of Personality and Social Psychology*, *84*(5), 960–971.

Anderson, C. A., Shibuya, A., Ihori, N., Swing, E. L., Bushman, B. J., Sakamoto, A., et al. (2010). Violent video game effects on aggression, empathy, and prosocial behavior in eastern and western countries: A meta-analytic review. *Psychological Bulletin*, *136*(2), 151–173.

Appel, M. (2011). The effects of media violence. In D. Christie (Ed.), *Encyclopedia of peace psychology*. New Jersey: Wiley-Blackwell.

Appelbaum, R. (2017). *The aesthetics of violence: Art, fiction, drama and film*. London: Rowman & Littlefield.

Archard, D. (2014). Insults, free speech and offensiveness. *Journal of Applied Philosophy*, *3*(2), 127–141.

Aristotle. (1999). *Nicomachean ethics* (T. Irwin, trans., 2nd Edition). Indianapolis: Hackett.

Aristotle. (2013). *Poetics* (A. Kenny, trans.). Oxford: Oxford University Press.

Ashton, S., McDonald, K., & Kirkman, M. (2019). What does "pornography" mean in the digital age? Revisiting a definition for social science researchers. *Porn Studies*, *6*(2), 144–168.

Atkin, C. (1983). Effects of realistic TV violence vs fictional violence on aggression. *Journalism and Mass Communication Quarterly*, *60*(4), 615–621.

Austin, J. L. (1962). *How to do things with words*. Oxford: Clarendon Press.

Ayer, A. J. (1953). Cogito, ergo sum. *Analysis*, *14*(2), 27–31.

Bacon, H. (2015). *The fascination of film violence*. London: Palgrave Macmillan.

Barongan, C., & Nagayama Hall, G. C. (1995). The influence of misogynous rap music on sexual aggression against women. *Psychology of Women Quarterly*, *19*, 195–207.

Barranco, R. E., Rader, N. E., & Smith, A. (2017). Violence and the box office: Considering ratings, ticket sales, and content. *Communications Research*, *44*(1), 77–95.

Bartel, C. (2012). Resolving the gamer's dilemma. *Ethics and Information Technology*, *14*(1), 11–16.

Bartel, C. (2015). Free will and moral responsibility in video games. *Ethics and Information Technology*, *17*(4), 285–293.

Bartel, C., & Cremaldi, A. (2018). "It's just a story": Pornography, desire, and the ethics of fictive imagining. *British Journal of Aesthetics*, *58*(1), 37–50.

Bell, J. (2008). Behind this mortal bone: The (in)effectiveness of torture. *Indiana Law Journal*, *83*(1), 339–361.

Bensley, L., & Van Eenwyk, J. (2001). Video games and real-life aggression: Review of the literature. *Journal of Adolescent Health*, *29*, 244–257.

Berlin, F. S., & Sawyer, D. (2012). Potential consequences of accessing child pornography over the internet and who is accessing it. *Sexual Addiction & Compulsivity: The Journal of Treatment & Prevention*, *19*(1–2), 30–40.

Biber, K. (2009). Bad holocaust art. *Law Text Culture*, *13*(1), 226–259.

Bird, P. (2011). Virtual child pornography and the constraints imposed by the first amendment. *Barry Law Review*, *16*(1), 161–176.

Bivona, J., & Critelli, J. (2009). The nature of women's rape fantasies: An analysis of prevalence, frequency, and contents. *Journal of Sex Research*, *46*(1), 33–45.

Björnsson, G., & McPherson, T. (2014). Moral attitudes for non-cognitivists: Solving the specification problem. *Mind*, *489*(1), 1–38.

Blackburn, S. (1985). Supervenience revisited. In I. Hacking (Ed.), *Exercises in analysis: Essays in honour of Casimir Lewy* (pp. 47–67). Cambridge: Cambridge University Press.

Blackburn, S. (2009). Truth and a priori possibility: Egan's charge against quasi-realism. *Australasian Journal of Philosophy*, *87*(2), 201–213.

Booth, W. C. (1964). Censorship and the values of fiction. *The English Journal*, *53*(3), 155–164.

Booth, W. C. (1988). *The company we keep: An ethics of fiction*. Berkeley: University of California Press.

Borowitz, A. I. (1973). Salieri and the "murder" of Mozart. *The Musical Quarterly*, *59*(2), 263–284.

Bourke, V. J. (1956). Moral problems related to censoring the media of mass communications. *Marquette Law Review*, *40*(1), 57–73.

Bradley, B. (2012). Doing away with harm. *Philosophy and Phenomenological Research*, *85*(2), 390–412.

Brey, P. (2003). The social ontology of virtual environments. *American Journal of Economics and Sociology*, *62*(1), 269–281.

Brock, S. (2010). The creationist fiction: The case against creationism about fictional characters. *Philosophical Review*, *119*(3), 337–364.

Brodsky, W., Olivieri, D., & Chekaluk, E. (2018). Music genre induced driver aggression: A case of music delinquency and risk promoting popular culture. *Music and Science*, *1*, 1–17.

Brown, J., & St Clair, N. (2006). *The distant mirror: Reflections on young adult historical fiction*. Lanham, MD: Scarecrow Press.

Brown, M. (2009). The shocking Victorian murder that showed Charles Dickens how to finish off Nancy. *The Guardian*, 11 April. Retrieved 21 February 2019 from https://www.theguardian.com/books/2009/apr/11/dickens-oliver-twist-eliza-grimwood-murder.

Brown, P. (1992). Amadeus and Mozart: Setting the record straight. *The American Scholar*, *61*(1), 49–66.

Brown, R. S. (2007). Torture, terrorism, and the ticking bomb: A principled response. *Journal of International Law & Policy*, *4*, 1–33.

Browne, K. D., & Hamilton-Giachritisis, C. (2005). The influence of violent media on children and adolescents: A public health approach. *The Lancet*, *365*(9460), 702–710.

Brummert Lennings, H. I., & Warburton, W. A. (2011). The effect of auditory versus visual violent media exposure on aggressive behaviour: The role of song lyrics, video clips and musical tone. *Journal of Experimental Social Psychology*, *47*, 794–799.

Bufacchi, V., & Arrigo, J. M. (2006). Torture, terrorism and the state: A refutation of the ticking-bomb argument. *Journal of Applied Philosophy*, *23*(3), 355–373.

Busching, R., Allen, J. J., & Anderson, C. A. (2016). Violent media content and effects. In J. F. Nussbaum (Ed.), *Oxford research encyclopedia of communication* (pp. 1–21). New York: Oxford University Press.

Bushman, B. J., Gollwitzer, M., & Cruz, C. (2015). There is broad consensus: Media researchers agree that violent media increase aggression in children, and pediatricians and parents concur. *Psychology of Popular Media Culture*, *4*(3), 200–214.

Bushman, B. J., & Huesmann, L. R. (2014). Twenty-five years of research on violence in digital games and aggression revisited: A reply to Elson and Ferguson (2013). *European Psychologist*, *19*(1), 47–55.

Bushman, B. J., Jamieson, P. E., Weitz, I., & Romer, D. (2013). Gun violence trends in movies. *Pediatrics*, *132*(6), 1014–1018.

Bushman, B. J., Rothstein, H. R., & Anderson, C. A. (2010). Much ado about something: Violent video game effects and a school of red herring: Reply to Ferguson and Kilburn (2010). *Psychological Bulletin*, *136*(2), 182–187.

Button, T. (2012). Spotty scope and our relation to fiction. *Noûs*, *46*(2), 243–258.

Cahn, S. M. (1974). A clockwork orange is not about violence. *Metaphilosophy*, *5*(2), 155–157.

Calvert, S. L., Appelbaum, M., Dodge, K. A., Graham, S., Nagayama Hall, G. C., Hamby, S., et al. (2017). The American Psychological Association Task Force assessment of violent video games: Science in the service of public interest. *American Psychologist*, *72*(2), 126–143.

Carlson, E. (2019). More problems for the counterfactual comparison account of harm and benefit. *Ethical Theory and Moral Practice*, *22*(4), 795–807.

Carroll, N. (1990). *The philosophy of horror, or, paradoxes of the heart*. New York: Routledge.

Carroll, N. (2002). The wheel of virtue: Art, literature, and moral knowledge. *The Journal of Aesthetics and Art Criticism*, *60*(1), 3–26.

Carroll, N. (2014). Ethics and comic amusement. *British Journal of Aesthetics*, *54*(2), 241–253.

Casey, D. (2000). Sacrifice, piss Christ, and liberal excess. *Law Text Culture*, *5*(1), 1–9.

Casey, M., Fischer, A., & Ramsey, H. (2000). After Serrano ethics, theology and the law of blasphemy. *Law Text Culture*, *5*(1), 1–11.

Cherry, C. (1988). When is fantasising morally bad? *Philosophical Investigations*, *11*(2), 112–132.

Chick, T. (2009). Is modern warfare 2 the most disgusting game of the year? *Fidgit*, 10 November 2009. In Young, G. (2014). *Ethics in the virtual world: The morality and psychology of gaming*. London: Routledge.

Chrisman, M. (2016). *The meaning of "ought". Beyond descriptivism and expressivism in metaethics*. Oxford: Oxford University Press.

Christen, M., & Katsarov, J. (2016). Moral sensitivity and preconditions of moral distress. *The American Journal of Bioethics*, *16*(2), 19–21.

CNN. (2003). "Master" and "slave" computer labels unacceptable, officials say. *CNN.com*, 26 November. Retrieved 22 July 2019 from http://edition.cnn.com/2003/TECH/ptech/11/26/master.term.reut/.

Cooke, B. (2014). Ethics and fictive imagining. *The Journal of Aesthetics and Art Criticism*, *72*(3), 317–327.

Coons, C. (2011). How to prove that some acts are wrong (without using substantive moral premises). *Philosophical Studies*, *155*, 83–98.

Corvino, J. (2002). Naughty fantasies. *Southwest Philosophy Review*, *18*(1), 213–220.

Coyne, S. M., Callister, M. A., Gentile, D. A., & Howard, E. (2016). Media violence and judgments of offensiveness: A quantitative and qualitative analysis. *Psychology of Popular Media Culture*, *5*(4), 372–389.

Coyne, S. M., Callister, M., Pruett, T., Nelson, D. A., Stockdale, L., & Wells, B. M. (2011). A mean read: Aggression in adolescent English literature. *Journal of Children and Media*, *5*(4), 411–425.

Coyne, S. M., Callister, M., Stockdale, L., Coutts, H., & Collier, K. M. (2015). Just how graphic are graphic novels? An examination of aggressive behaviour in adolescents. *Violence and Victims*, *30*(2), 208–224.

Coyne, S. M., Ridge, R., Stevens, M., Callister, M., & Stockdale, L. (2012). Backbiting and bloodshed in books: Short-term effects of reading physical and relational aggression in literature. *British Journal of Social Psychology*, *51*, 188–196.

Critelli, J. W., & Bivona, J. M. (2008). Women's erotic rape fantasies: An evaluation of theory and research. *Journal of Sex Research*, *45*, 57–70.

Cunningham, H. (2014). Did Salieri really kill Mozart? *Sydney Morning Herald*, 10 June. Retrieved 15 November 2019 from https://www.smh.com.au/entertainment/opera/did-salieri-really-kill-mozart-20140612-zs2t4.html.

Cunningham, S., Engelstätter, B., & Ward, M. R. (2016). Violent video games and violent crime. *Southern Economic Journal*, *82*(4), 1247–1265.

Currie, G. (1990). *The nature of fiction*. Cambridge: Cambridge University Press.

Currie, G. (1995). The moral psychology of fiction. *Australasian Journal of Philosophy*, *73*(2), 250–259.

Curtis, C. (n.d.) Deepfakes are being weaponized to silence women but this woman is fighting back. Retrieved 11 October 2019 from https://thenextweb.com/code-word/2018/10/05/deepfakes-are-being-weaponized-to-silence-women-but-this-woman-is-fighting-back/.

Cushman, F., Gray, K., Gaffey, A., & Mendes, W. B. (2012). Simulating murder: The aversion to harmful action. *Emotion*, *12*(1), 2–7.

Daniels, C. B., & Scully, S. (1992). Pity, fear and catharsis in Aristotle's poetics. *Noûs*, *26*(2), 204–217.

Darwall, S. L. (1977). Two kinds of respect. *Ethics*, *88*(1), 36–49.

Davies, D. (1996). Fictional truth and fictional authors. *British Journal of Aesthetics*, *36*, 43–55.

Davies, D. (2001). Fiction. In B. Gaut & D. McIver Lopes (Eds.), *The Routledge companion to aesthetics* (pp. 263–274). London: Routledge.

Davies, P. J. (1983). Mozart's illnesses and death. *Journal of the Royal Society of Medicine*, *76*, 776–785.

de Groot, J. (2010). *The historical novel*. London: Routledge.

DeCamp, W., & Ferguson, C. J. (2017). The impact of degree of exposure to violent video games, family background, and other factors on youth violence. *Journal of Youth and Adolescence*, *46*, 388–400.

Delfino, R. A. (2019). Pornographic deepfakes—Revenge porn's next tragic act: The case for federal criminalization. *Fordham Law Review*, *88*(3), article 2, 887–938.

Demos, J. (2005). Afterword: Notes from, and about, the history/fiction borderland. *Rethinking History*, *9*(2–3), 329–335.

Depaul, M. R. (1988). Arguments and perception: The role of literature in moral inquiry. *The Journal of philosophy*, *85*(10), 552–565.

Depaul, R. (1987). Supervenience and moral dependence. *Philosophical Studies*, *51*(3), 425–439.

Dershowitz, A. M. (2004). Tortured reasoning. In S. Levinson (Ed.), *Torture: A collection* (pp. 257–280). Oxford: Oxford University Press.

Devereaux, M. (2004). Moral judgment and works of art: The case of narrative literature. *The Journal of Aesthetic and Art Criticism*, *62*(1), 3–11.

Di Muzio, G. (2006). The immorality of horror films. *International Journal of Applied Philosophy*, *20*, 277–294.

Dillon, M. C. (1982). The phenomenon of obscenity in literature: The specification of a value. *Journal of Value Inquiry*, *16*(4), 259–274.

Driver, J. (1992). The suberogatory. *Australasian Journal of Philosophy*, *70*(3), 286–295.

Dupovy-Camet, J. (2002). Trichinellosis is unlikely to be responsible for Mozart's death. *Archives of Internal Medicine*, *162*(8), 946–947.

Durkheim, E. (1982). *The Rules of Sociological Method* (S. Lukes, trans.). New York: Free Press. (Originally published 1895).

Dworkin, R. (2006). A new map of censorship. *Index on Censorship*, *35*(1), 130–133.

Eagle, A. (2007). Telling tales. *Proceedings of the Aristotelian Society*, *107*(2), 125–147.

Edelstein, D. (2006). Now playing at your local multiplex: Torture porn. *NewYork Movies*, 28 January. Retrieved 10 December 2019 from http://nymag.com/movies/features/15622/.

Eden, K. (1982). Poetry and equity: Aristotle's defense of fiction. *Traditio*, *38*, 17–43.

Edwards, S. S. M. (2000). Prosecuting "child pornography": Possession and taking of indecent photos of children. *Journal of Social Welfare and Family Law*, *22*, 1–21.

Eidelson, B. (2015). *Discrimination and disrespect*. Oxford: Oxford University Press.

Eneman, M., Gillespie, A. A., & Stahl, B. C. (2009). *Criminalising fantasies: The regulation of virtual child pornography* (pp. 1–12). Retrieved 7 December 2019 from https://www.academia.edu/21221849/Criminalising_fantasies_The_regulation_of_virtual_child_pornography.

Evans, M. (2016). Police win sandcastle competition with "poor taste" sculpture of naked murder victim. *The Telegraph*, 20 May. Retrieved 19 August 2019 from http://www.telegraph.co.uk/news/2016/05/20/police-win-sandcastle-competition-with-poor-taste-sculpture-of-n/.

Everett, A. (2007). Pretense, existence, and fictional objects. *Philosophy and Phenomenological Research*, *74*(1), 56–80.

Fatsis, L. (2019). Policing the beats: The criminalisation of UK drill and grime music by the London metropolitan police. *The Sociological Review*, *67*(6), 1300–1316.

Feagin, S. L. (2010). Film appreciation and moral insensitivity. *Midwest Studies in Philosophy*, *43*, 20–33.

Feinberg, J. (1984). *Harm to others: The moral limits of criminal law, volume 1*. Oxford: Oxford University Press.

Feinberg, J. (1988). *The moral limits of criminal law, volume 2: Offense to others*. Oxford: Oxford University Press.

Feit, N. (2015). Plural harm. *Philosophy and Phenomenological Research*, *90*(2), 361–388.

Feit, N. (2019). Harming by failing to benefit. *Ethical Theory and Moral Practice*, *22*(4), 809–823.

Feldman, G., Lian, H., Kosinski, M., & Stillwell, D. (2017). Frankly, we do give a damn: The relationship between profanity and honesty. *Social Psychology and Personality Science*, *8*(7), 816–826.

Ferguson, C. J. (2007a). Evidence for publication bias in video game violence effects literature: A meta-analytic review. *Aggression and Violent Behavior*, *12*, 470–482.

Ferguson, C. J. (2007b). The good, the bad and the ugly: A metaanalytic review of positive and negative effects of violent video games. *Psychiatric Quarterly*, *78*(4), 309–316.

Ferguson, C. J. (2008). The school shooting/violent video game link: Causal relationship or moral panic? *Journal of Investigative Psychology and Offender Profiling*, *5*, 25–37.

Ferguson, C. J. (2009). Media violence effects: Confirmed truth or just another x-file? *Journal of Forensic Psychology Practice*, *9*(2), 103–126.

Ferguson, C. J. (2011). Video games and youth violence: A prospective analysis in adolescents. *Journal of Youth and Adolescence*, *40*, 377–391.

Ferguson, C. J. (2013). Violent video games and the supreme court: Lessons for the scientific community in the wake of Brown v. Entertainment Merchants Association. *American Psychologist*, *68*(2), 57–74.

Ferguson, C. J. (2015). Does movie or video game violence predict societal violence? *Journal of Communication*, *61*(1), E1–E22.

Ferguson, C. J., & Kilburn, J. (2009). The public health risks of media violence: A meta-analytic review. *Journal of Paediatrics*, *154*(5), 759–763.

Ferguson, C. J., & Kilburn, J. (2010). Much ado about nothing: The misestimation and overinterpretation of violent video game effects in eastern and western nations: Comment on Anderson et al. (2010). *Psychological Bulletin*, *136*(2), 174–178.

Fischer, P., & Greitemeyer, T. (2006). Music and aggression: The impact of sexual aggressive son lyrics on aggression-related thoughts, emotion, and behaviour towards the same or opposite sex. *Personality and Social Psychology Bulletin*, *3*(9), 1165–1176.

Foot, P. (1967). The problem of abortion and the doctrine of double effect. *Oxford Review*, *5*, 1–5.

Freedman, J. (1984). Effect of television violence on aggressiveness. *Psychological Bulletin*, *96*(2), 227–245.

Freedman, J. (2002). *Media violence and its effect on aggression: Assessing the scientific evidence*. Toronto: University of Toronto Press.

Friend, S. (2007). Fictional characters. *Philosophy Compass*, *2*(2), 141–156.

Friend, S. (2012). Fiction as a genre. *Proceedings of the Aristotelian Society*, *62*(2), 179–209.

Funk, N. (1984). Mill on censorship. *History of Philosophy Quarterly*, *1*(4), 453–463.

Garrison, D. (1976). Immoral fiction in the late Victorian Library. *American Quarterly*, *28*(1), 71–89.

Gaut, B. (1998). The ethical criticism of art. In J. Levinson (Ed.), *Aesthetics and ethics: Essays at the intersection* (pp. 182–203). Cambridge: Cambridge University Press.

Gaut, B. (2003). Creativity and imagination. In B. Gaut & P. Livingston (Eds.), *The creation of art: New essays in philosophical aesthetics* (pp. 148–173). Cambridge: Cambridge University Press.

Gaut, B. (2007). *Art, emotion and ethics*. Oxford: Oxford University Press.

Gendler, T. S. (2000). The puzzle of imaginative resistance. *Journal of Philosophy*, *97*(2), 55–81.

Gendler, T. S. (2006). Imaginative resistance revisited. In S. Nichols (Ed.), *The architecture of the imagination: New essays on pretence, possibility, and fiction* (pp. 149–175). Oxford: Oxford University Press.

Gentile, D. A. (2013). Catharsis and media violence: A conceptual analysis. *Societies*, *3*, 491–510.

Gert, B. (1969). Justifying violence. *The Journal of Philosophy*, *66*(19), 616–628.

Gert, B. (2004). *Common morality: Deciding what to do*. Oxford: Oxford University Press.

Gert, B. (2011). The usefulness of a comprehensive and systematic moral theory. *Teaching Ethics*, *12*(1), 25–38.

Gethering, A. (2012). In defence of fiction: History and imagination in Kate Grenville's *the secret river* and *the lieutenant*. In K. Cooper & E. Short (Eds.), *The female figure in contemporary history fiction* (pp. 189–205). London: Palgrave Macmillan.

Gillespie, A. (2010). Legal definitions of child pornography. *Journal of Sexual Aggression*, *16*(1), 19–31.

Goerger, M. (2017). Value, violence, and the ethics of gaming. *Ethics and Information Technology*, *19*(2), 95–105.

Goodwin, G. P., & Darley, J. M. (2012). Why are some moral beliefs perceived to be more objective than others? *Journal of Experimental Social Psychology*, *48*(1), 250–256.

Goranson, R. E. (1970). Media violence and aggressive behaviour: A review of experimental research. *Advances in Experimental Social Psychology*, *5*, 1–31.

Greitemeyer, T. (2011). Exposure to music with prosocial lyrics reduces aggression: First evidence and test of underlying mechanism. *Journal of Experimental Social Psychology*, *47*, 28–36.

Greitemeyer, T., & Mügge, O. (2014). Video games do affect social outcomes: A meta-analytic review of the effects of violent and prosocial video game play. *Personality and Social Psychology Bulletin*, *40*(5), 578–589.

Greitemeyer, T., & Sagioglou, C. (2017). The longitudinal relationship between everyday sadism and the amount of violent video game play. *Personality and Individual Differences*, *104*, 238–242.

Grimshaw, J. (1993). Ethics, fantasy, and self-transformation. *Royal Institute of Philosophy Supplement*, *35*, 145–158.

Grittenden, C. (1966). Fictional existence. *American Philosophical Quarterly*, *3*(4), 317–321.

Groce, E., & Groce, R. (2005). Authenticating historical fiction: Rationale and process. *Educational Research Perspectives*, *32*(1), 99–119.

Gubler, J. R., Herrick, S., Price, R. A., & Wood, D. A. (2018). Violence, aggression, and ethics: The link between exposure to human violence and unethical behaviour. *Journal of Business Ethics*, *147*, 25–35.

Güera, D., & Delp, E. J. (2018). Deepfake video detection using recurrent neural networks. In *IEEE international conference on advanced video and signal based surveillance (AVSS)* (pp. 1–6), Piscataway, NJ: IEEE.

Gunter, B. (2008). Media violence: Is there a case for causality? *American Behavioral Scientist*, *51*(8), 1061–1122.

Hanser, M. (2008). The metaphysics of harm. *Philosophy and Phenomenological Research*, *77*(2), 421–450.

Hanser, M. (2011). Still more on the metaphysics of harm. *Philosophy and Phenomenological Research*, *82*(2), 459–469.

Hanser, M. (2019). Understanding harm and its moral significance. *Ethical Theory and Moral Practice*, *22*(4), 853–870.

Hare, R. M. (1952). *The language of morals*. Oxford: Clarendon Press.

Harman, G., & Jarvis Thomson, J. (1996). *Moral relativism and moral objectivity*. Oxford: Blackwell.

Harold, J. (2003). Flexing the imagination. *The Journal of Aesthetics and Art Criticism*, *61*(3), 247–257.

Harris, D. (2019). Deepfakes: False pornography is here and the law cannot protect you. *Duke Law & Technology Review*, *17*, 99–128.

Harris, J. (1998). *Clones, genes and immortality*. Oxford: Oxford University Press.

Harrosh, S. (2012). Identifying harms. *Bioethics*, *26*(9), 493–498.

Harwell, D. (2019). Scarlett Johansson on fake AI-generated sex videos: "Nothing can stop someone from cutting and pasting my image." *The Washington Post*, 1 January. Retrieved 11 October 2019 from https://www.washingtonpost.com/technology/2018/12/31/scarlett-johansson-fake-ai-generated-sex-videos-nothing-can-stop-someone-cutting-pasting-my-image/?noredirect=on.

Haynes, N. (2016). Rebooting the blockbuster, who ya gonna call? Aristotle, actually. *The Guardian*, 11 August. Accessed 2 July 2019 from https://www.theguardian.com/commentisfree/2016/aug/10/summer-blockbuster-aristotle-suicide-squad-bourne-star-trek-exercise-peppa-pig.

Henry, N., Flynn, A., & Powell, A. (2018). AI porn can now create fake porn, making revenge porn even more complicated. *The Conversation*, 1 March. Retrieved 10 October 2019 from https://theconversation.com/ai-can-now-create-fake-porn-making-revenge-porn-even-more-complicated-92267.

Henry, N., Flynn, A., & Powell, A. (2019). *Responding to revenge pornography: Prevalence, nature and impacts*. Canberra: Australian Institute of Criminology.

Hershfield, J. (2009). The ethics of sexual fantasy. *International Journal of Applied Philosophy*, *23*(1), 27–49.

Hill, J. M. (1987). Pornography and degradation. *Hypatia*, *2*(2), 39–54.

Hinte, F. (1971). Should desensitization techniques be part of police recruit training? *Group Psychotherapy and Psychodrama*, *24*(3–4), 107–110.

Hintikka, J. (1962). Cogito, ergo sum: Inference or performance? *The Philosophical Review*, *71*(1), 3–32.

Hirschmann, J. V. (2001). What killed Mozart? *Archives of Internal Medicine*, *161*(11), 1381–1389.

Holpuch, A. (2012). Andres Serrano's controversial Piss Christ goes on view in New York. *The Guardian*, 29 September. Retrieved 9 December 2019 from https://www.theguardian.com/artanddesign/2012/sep/28/andres-serrano-piss-christ-new-york.

Honneth, A. (1992). Integrity and disrespect: Principles of a conception of morality based on the theory of recognition. *Political Theory*, *20*(2), 187–201.

Hopkins, P. (1994). Rethinking sadomasochism: Feminism, interpretation, and simulation. *Hypatia*, *9*(1), 116–141.

Howitt, D. (1995). Pornography and the paedophile: Is it criminogenic? *British Journal of Medical Psychology*, *68*, 15–27.

Huesmann, L. R. (2007). The impact of electronic media violence: Scientific theory and research. *Journal of Adolescent Health*, *41*(6 Supplement 1), S6–S13.

Huesmann, L. R. (2010). Nailing the coffin shut on doubts that violent video games stimulate aggression: Comment on Anderson et al. (2010). *Psychological Bulletin*, *136*(2), 179–181.

Huesmann, L. R., Dubow, E. F., & Yang, G. (2013). Why it is hard to believe that media violence causes aggression? In K. E. Dill (Ed.), *The Oxford handbook of media psychology* (pp. 159–171). New York: Oxford University Press.

Huesmann, L. R., & Taylor, L. D. (2006). The role of media violence in violent behaviour. *Annual Review of Public Health*, *27*, 393–415.

Hume, D. [1757] (1995). Four dissertations. London: Thoemmes Continuum.

Hunt, M. (1974). *Sexual behavior in the 1970's*. New York: Playboy Press.

Hunter, I., Saunders, D., & Williamson, D. (1993). *On pornography: Literature, sexuality and obscenity law*. London: Macmillan.

Irving, H., & Townend, J. (2016). Censorship and national security: Information control in the second world war and present day. *History and Policy*. Retrieved 23 January 2020 from http://www.historyandpolicy.org/policy-papers/papers/censorship-and-national-security-information-control.

Jackson, F., & Pettit, P. (1998). A problem for expressivism. *Analysis*, *58*(4), 239–251.

James, D. N. (1993). The ethics of fantasising. *International Journal of Applied Philosophy*, *8*(1), 51–55.

Jansz, J. (2005). The emotional appeal of violent video games for adolescent males. *Communication Theory*, *15*(3), 219–241.

Jay, T., & Janschewitz, K. (2008). The pragmatics of swearing. *Journal of Politeness Research*, *4*, 267–288.

Jenkins, P. (2001). *Beyond tolerance: Child pornography on the internet*. New York: New York University Press.

Johnston, D. D. (1995). Adolescents' motivations for viewing graphic horror. *Human Communication Research*, *21*(4), 522–552.

Joyal, C. C., Amélie Cossette, A., & Lapierre, V. (2015). What exactly is an unusual sexual fantasy? *Journal of Sexual Medicine*, *12*(2), 328–340.

Kahane, G., & Savulescu, J. (2012). The concept of harm and the significance of normality. *Journal of Applied Philosophy*, *29*(4), 318–332.

Kajtár, L. (2017). Fiction cannot be true. *Philosophical Studies*, *174*, 2167–2186.

Kalim, A. (2013). Addressing the gap in international instruments governing internet child pornography. *CommLaw Conspectus*, *21*, 428–452.

Kant, I. [1785] (1993). *Grounding for the metaphysics of morals* (J. W. Ellington, trans.). Indianapolis: Hackett Publishing Company, Inc.

Karhausen, L. R. (2010). Mozart's 140 causes of death and 27 mental disorders. *British Medical Journal*, *341*, c6789.

Karnold, S. (2000). The cyber world of child pornography and the Child Pornography Act of 1996: Thoughts on morphing, virtual imaging and the first amendment. *Journal of Information Ethics*, *2*, 60–65.

Keefe, S. P. (2009). Beyond fact and fiction, scholarly and popular: Peter Shaffer and Miloš Forman's "Amadeus" at 25. *The Musical Times*, *150*(1906), 45–53.

Kershnar, S. (2005). The moral status of sexual fantasies. *Public Affairs Quarterly*, *19*(4), 301–315.

Khalid, A. (2019). Deepfake videos are a far, far bigger problem for women. *Quartz*, 9 October. Retrieved 10 October 2019 from https://qz.com/1723476/deepfake-videos-feature-mostly-porn-according-to-new-study-from-deeptrace-labs/.

Kieran, M. (1996). Art, imagination, and the cultivation of morals. *The Journal of Aesthetics and Art Criticism*, *54*(4), 337–351.

Kieran, M. (2002). On obscenity: The thrill and repulsion of the morally prohibited. *Philosophy and Phenomenological Research*, *64*(1), 31–55.

Kim, J. (1993). Supervenience and mind: Selected philosophical essays. Cambridge: Cambridge University Press.

Kingston, D. A., Fedoroff, P., Firestone, P., Curry, S., & Bradford, J. M. (2008). Pornography use and sexual aggression: The impact of frequency and type of pornography use on recidivism among sexual offenders. *Aggressive Behavior*, *34*(4), 341–351.

Kirsh, S. J. (2006). Cartoon violence and aggression in youth. *Aggression and Violent Behavior*, *11*, 547–557.

Kirsh, S. J., & Olczak, P. V. (2000). Violent comic books and perceptions of ambiguous provocation situations. *Media Psychology*, *2*, 47–62.

Kirsh, S. J., & Olczak, P. V. (2002a). Violent comic books and judgments of relational aggression. *Violence and Victims*, *17*, 373–380.

Kirsh, S. J., & Olczak, P. V. (2002b). The effects of extremely violent comic books on social information processing. *Journal of Interpersonal Violence*, *17*(11), 1160–1178.

Kjeldgaard-Christiansen, J. (2020). Splintering the gamer's dilemma: Moral intuitions, motivational assumptions, and action prototypes. *Ethics and Information Technology*, *22*(1), 93–102.

Köhler, S. (2013). Do expressivists have an attitude problem? *Ethics*, *123*(3), 479–507.

Konijn, E. A., & Hoorn, J. F. (2005). Some like it bad: Testing a model for perceiving and experiencing fictional characters. *Media Psychology*, *7*(2), 107–144.

Koppelman, A. (2005). Does obscenity cause moral harm? *Columbia Law Review*, *105*, 1635–1680.

Kosse, S. H. (2004). Virtual child pornography—A United States update. *Communications Law*, *9*(2), 39–46.

Krahé, B. (2014). Restoring the spirit of fair play in the debate about violent video games: A comment on Elson and Ferguson (2013). *European Psychologist*, *19*(1), 56–59.

Kreider, S. E. (2008). The virtue of horror films: A response to Di Muzio. *International Journal of Applied Philosophy*, *22*(1), 149–157.

Kripke, S. (1980). *Naming and necessity*. Cambridge, MA: Harvard University Press.

Lamarque, P. (2010). *Works and objects*. Oxford: Oxford University Press.

Lamont, C. (2018). Killing time. *The Scottish Sun*, 7 February. Retrieved 3 July 2019 from https://www.thescottishsun.co.uk/tvandshowbiz/tv/2196981/actor-douglas-henshall-tv-show-shetland-returns-island-murder-capital/.

Langbein, J. H. (2004). The legal history of torture. In S. Levinson (Ed.), *Torture: A collection* (pp. 93–104). Oxford: Oxford University Press.

Lange, J. (2019). In defense of Quentin Tarantino's over-the-top violence. *The Week*, 25 July. Retrieved 17 December 2019 from https://theweek.com/articles/854623/defense-quentin-tarantinos-overthetop-violence.

Laumann, E., Gagnon, J., Michael, R., & Michaels, S. (1994). *The social organization of sexuality*. Chicago: University of Chicago Press.

Leclerc, A. (2016). Actualism and fictional characters. *Principia*, *20*(1), 61–80.

Lee, D. (2018). Deepfakes porn. Has serious consequences. *BBC News*, 3 February. Retrieved 10 October 2019 from https://www.bbc.com/news/technology-42912529.

Left, S. (2002). White supremacists create racist computer games. *The Guardian*, 21 February. Retrieved 12 December 2018 from https://www.theguardian.com/technology/2002/feb/21/games.internetnews.

Levy, N. (2002). Virtual child pornography: The eroticization of inequality. *Ethics in Information Technology*, *4*, 319–323.

Lewis, D. (1986). *On the plurality of worlds*. Oxford: Basil Blackwell.

Liao, S-Y. (2013). Moral persuasion and the diversity of fictions. *Pacific Philosophical Quarterly*, *94*, 269–289.

Liao, S.-Y., & Gendler, T. S. (2015). The problem of imaginative resistance. In N. Carroll & J. Gibson (Eds.), *The Routledge companion to the philosophy of literature* (pp. 405–418). London: Routledge.

Lindsay, J. J., Dill-Shackleford, K. E., Anderson, K. B., & Bartholow, B. D. (2014). The proliferation of media violence and its economic underpinnings. In D. A. Gentile (Ed.), *Media violence and children* (2nd ed., pp. 1–44). Santa Barbara, CL: Praeger.

Livingston, P., & Sauchelli, A. (2011). Philosophical perspectives on fictional characters. *New Literary History*, *42*(2), 337–360.

Lovett, B., & Jordan, A. H. (2010). Levels of moralisation: A new conception of moral sensitivity. *Journal of Moral Education*, *39*(2), 175–189.

Lowth, M. (2017). Does torture work? Donald Trump and the CIA. *British Journal of General Practice*, *67*(656), 126.

Luck, M. (2009). The gamer's dilemma: An analysis of the arguments for the moral distinction between virtual murder and virtual paedophilia. *Ethics and Information Technology*, *11*, 31–36.

Ludlow, P. (2006). From Sherlock and Buffy to Klingon and Norrathian Platinum Pieces: Pretense, contextualism, and the myth of fiction. In E. Sosa & E. Villanueva (Eds.), *Philosophical issues, vol. 16: Philosophy of language* (pp. 162–183). Oxford: Blackwell.

Lukács, G. (1983). *The historical novel* (H. Mitchell & S. Mitchell, trans.). Lincoln: University of Nebraska Press.

MacDonald, C., & Palmer, A. (2019). Horrifying deepfake video blends Jennifer Lawrence and Steve Buscemi in latest example of the "nightmare" technology. *The Daily Mail Australia*, 2 February. Retrieved 11 October 2019 from https://www.dailymail.co.uk/sciencetech/article-6655231/Horrifying-deepfake-video-blends-Jennifer-Lawrence-Steve-Buscemi.html.

Machkovech, S. (2015). After outcry, "edutainment" game removes slave-Tetris mode. Ars Technica. Retrieved 7 December 2019 from http://arstechnica.com/gaming/2015/09/after-outcry-edutainment-game-removes-slave-tetris-mode/.

Malliet, S. (2006). An exploration of adolescents' perceptions of videogame realism. *Learning, Media and Technology*, *31*(4), 377–394.

Markey, P. M., Markey, C. N., & French, J. E. (2015). Violent video games and real-world violence: Rhetoric versus data. *Psychology of Popular Media Culture*, *4*(4), 277–295.

Mast, J. F., & McAndrew, F. T. (2011). Violent lyrics in heavy metal music can increase aggression in males. *North American Journal of Psychology*, *13*(1), 63–64.

May, L. (1992). Insensitivity and moral responsibility. *The Journal of Value Inquiry*, *26*, 7–22.

McAllister, M. (2006). Irreversible movie review. *Future Movies*, 13 December. Retrieved 7 December 2019 from http://www.futuremovies.co.uk/review.asp?ID=639.

McAteer, J. (2016). How to be a moral taste theorist. *Essays in Philosophy*, *17*(1), 5–21.

McCormick, K. (1977). Censorship: Some philosophical issues. *Index on Censorship*, *6*(2), 31–37.

McCormick, M. (2001). Is it wrong to play violent video games? *Ethics and Information Technology*, *3*(4), 277–287.

McGarry, D. D., & Harriman White, S. (1963). *Historical fiction guide: Annotated chronological, geographical and topical list of five thousand selected historical novels*. New York: The Scarecrow Press.

McGinn, C. (1997). *Ethics, evil and fiction*. Oxford: Oxford University Press.

McGlynn, C., & Rackley, E. (2009). Criminalising extreme pornography: A lost opportunity. *Criminal Law Review*, *4*, 245–260.

McMahan, J. (2006). On the moral equality of combatants. *Journal of Political Philosophy*, *14*(4), 377–393.

Meinong, A. (1960). The theory of objects (A. Levi, D. B. Terrell, & R. M. Chisholm, trans.). In R. M. Chisholm (Ed.), *Realism and the background of phenomenology* (pp. 76–117). New York: Free Press. (Originally published 1904).

Melville, K. (2019). The insidious rise of deepfake porn videos and one woman who won't be silenced. *ABC News*, 30 August. Retrieved 7 October 2019 from https://www.abc.net.au/news/2019-08-30/deepfake-revenge-porn-noelle-martin-story-of-image-based-abuse/11437774.

Merritt, S. (2014). How true should fiction be? *The Guardian*, 20 March. Retrieved 28 November 2019 from https://www.theguardian.com/books/booksblog/2014/mar/19/how-true-should-historical-fiction-be-mantel-andrew-miller-gregory.

Meyers, C. (2012). Expressivism, constructivism, and the supervenience of moral properties. *Philosophical Explorations*, *15*(1), 17–31.

Mill, J. S. (1957). Utilitarianism. In A. Piest (Ed.). New York: Bobbs-Merrill. (Originally published 1861).

Mill, J. S. (2005). *On liberty*. New York: Cosimo. (Originally published 1859).

Miller, A. (2003). *Introduction to contemporary metaethics*. Cambridge: Polity.

Miller, M. C. (1975). In defence of Sam Peckinpah. *Film Quarterly*, *28*, 3.

Miller, R., & Cushman, F. (2013). Aversive for me, wrong for you: First-person behavioral aversions underlie the moral condemnation of harm. *Social and Personality Psychology Compass*, *7*(10), 707–718.

Mills, C. (2000). Appropriating others' stories: Some questions about the ethics of writing fiction. *Journal of Social Philosophy*, *31*(2), 195–206.

Mitchell, E. (2003). Rape, violence... It's O.K. to look away. *The New York Times*, 7 March. Retrieved 7 December 2019 from http://movies.nytimes.com/movie/review?res=9E07EFDD133FF934A35750C0A9659C8B63.

Mole, C. (2009). Fiction's ontological commitments. *The Philosophical Forum*, *40*(4), 473–488.

Moore, G. E. (1903). *Principia ethica*. Cambridge: Cambridge university Press.

Moore, G. E. (1966). *Principia ethica*. Cambridge: Cambridge University Press.

Moore, N. (2013). Censorship is. *Australian Humanities Review*, *54*, 45–65.

Moran, R. (1994). The expression of feeling in imagination. *Philosophical Review*, *103*(1), 75–106.

Morris, I. (2018). Revenge "porn" gets even more horrifying with deepfakes. *Forbes*, 5 February. Retrieved 10 October 2019 from https://www.forbes.com/sites/ianmorris/2018/02/05/fakeapp-allows-anyone-to-make-deepfake-porn-of-anyone/#b52b40d391c9.

Mota, S. A. (2002). The U.S. supreme court addresses the child pornography prevention act and child online protection act in *Ashcroft v. Free Speech Coalition* and *Ashcroft v. American Civil Liberties Union. Federal Communications Law Journal*, *55*(1), 85–98.

Mothersill, M. (2006). Make-believe morality and fictional worlds. In J. L. Bermúdez & S. Gardner (Eds.), *Art and morality* (pp. 74–94). London: Routledge.

Mount, H. (2008). Deadly deeds in the city of bleeding spires. *The Telegraph*, 25 April. Retrieved 17 June 2019 from https://www.telegraph.co.uk/comment/personal-view/3557620/Deadly-deeds-in-the-city-of-bleeding-spires.html.

Mullin, A. (2004). Moral defects, aesthetic defects, and the imagination. *The Journal of Aesthetics and Art Criticism*, *62*(3), 249–261.

Nanay, B. (2010). Imaginative resistance and conversational implicature. *The Philosophical Quarterly*, *60*(240), 586–600.

Nauroth, P., Gollwitzer, M., Bender, J., & Rothmund, T. (2014). Gamers against science: The case of the violent video games debate. *European Journal of Social Psychology*, *44*, 104–116.

Neu, J. (2002). An ethics of fantasy? *Journal of Theoretical and Philosophical Psychology*, *22*(2) 133–157.

Neutze, J., Seto, M. C., Schaefer, G. A., Mundt, I. A., & Beier, K. M. (2011). Predictors of child pornography offenses and child sexual abuse in a community sample of pedophiles and hebephiles. *Sexual Abuse: A Journal of Research and Treatment*, *23*(2), 212–242.

Nichols, S. (2008). Sentimentalism naturalized. In W. Sinnott-Armstrong (Ed.), *Moral psychology: The evolution of morality* (Vol. 2, pp. 255–274). Cambridge, MA: MIT Press.

Nolan, D., & Sandgren, A. (2014). Creationism and cardinality. *Analysis*, *74*(4), 615–622.

Nolan, M. (2019). Novels are not instruction manuals, and Nobakov's Lolita is not an immoral book. *Newstatesman*, 27 March 2019. Retrieved 23 July 2020 from https://www.newstatesman.com/culture/books/2019/03/the-real-lolita-vladimir-nabokov-child-abuse-ficton-red-flag-sally-horner-sarah-weinman.

Nussbaum, M. C. (1992). *Love's knowledge. Essays on philosophy and literature*. Oxford: Oxford University Press.

Nys, T. (2010). Virtual ethics. *Ethical Perspectives*, *17*(1), 79–93.

O'Dea, C. J., Miller, S. S., Andres, E. B., Ray, M. H., Till, D. F., & Saucier, D. A. (2015). Out of bounds: Factors affecting the perceived offensiveness of racial slurs. *Language Sciences*, *52*, 155–164.

Öhman, C. (2020). Introducing the pervert's dilemma: A contribution to the criticism of Deepfake pornography. *Ethics and Information Technology*, *22*(2), 133–140.

O'Leary, C. (2016). Introduction: Censorship and creative freedom. In D. S. Sánchez & M. Thompson (Eds.), *Global insights on theatre censorship* (pp. 1–23). New York: Routledge.

Olson, L. (2013). *Those angry days*. New York: Random House.

Olulode, K. (2015). Airbrushing racism: Why racist words shouldn't be edited from history. *Index on Censorship*, *44*(4), 34–36.

O'Mara, S. (2015). *Why torture doesn't work: The neuroscience of interrogation*. Cambridge, MA: Harvard University Press.

Ostritsch, S. (2017). The amoralist challenge to gaming and the gamer's moral obligation. *Ethics and Information Technology*, *19*(2), 117–128.

Oxenbøll, M. (2017). Epistemologies of violence: Medieval Japanese war tales. *History and Theory*, *55*, 44–59.

Parson, T. (1980). *Non-existent objects*. New Haven: Yale University Press.

Patridge, S. (2011). The incorrigible social meaning of video game imagery. *Ethics and Information Technology*, *13*(4), 303–312.

Patridge, S. L. (2013). Pornography, ethics, and video games. *Ethics and Information Technology*, *15*(1), 25–34.

Peterson, S. B., & Lach, M. A. (1990). Gender stereotypes in children's books: Their prevalence and influence on cognitive and affective development. *Gender and Education*, *2*(2), 185–197.

Pitcher, G. (1984). The misfortunes of the dead. *American Philosophical Quarterly*, *21*, 183–188.

Plato. (2000). *The republic* (B. Jowett, trans.). New York: Dover Publications.

Pollock, J. L. (2006). *Thinking about acting: Logical foundations for rational decision making*. Oxford: Oxford University Press.

Poole, H. (1982). Obscenity and censorship. *Ethics*, *93*(1), 39–44.

Powell, R. (2010). What's the harm? An evolutionary theoretical critique of the precautionary principle. *Kennedy Institute of Ethics Journal*, *20*(2), 181–206.

Powers, T. M. (2003). Real wrongs in virtual communities. *Ethics in Information Technology*, *5*(4), 191–198.

Predelli, S. (1997). Talk about fiction. *Erkenntnis*, *46*, 69–77.

Prince, S. (1998). *Savage cinema: Sam Peckinpah and the rise of ultraviolent movies*. Austin: University of Texas Press.

Prinz, J. J. (2007). *The emotional construction of morals*. Oxford: Oxford University Press.

Przybylski, A. K., & Weinstein, N. (2019). Violent video game engagement is not associated with adolescents' aggressive behaviour: Evidence from a registered report. *Royal Society Open Science*, *6*, 171475, 1–16.

Purshouse, C. (2016). A defence of the counterfactual account of harm. *Bioethics*, *30*(4), 251–259.

Pushkin, A. (1933). Mozart and Salieri (A. F. B. Clark, trans.). *University of Toronto Quarterly*, *2*(4), 482–491. (Original publication 1830).

Quine, W. V. O. (1960). *Word and object*. Cambridge, MA: MIT Press.

Rabenberg, M. (2015). Harm. *Journal of Social Philosophy*, *8*(3), 1–33.

Ramirez, E. (2020). How to (dis)solve the Gamer's dilemma. Ethical Theory and Moral Practice, *23*(1), 141–161.

Rea, M. (2001). What is pornography? *Nous*, *35*(1), 118–145.

Ridge, M. (2006). Ecumenical expressivism: Finessing frege. *Ethics*, *116*(2), 302–336.

Ridge, M. (2007). Anti-reductionism and supervenience. *Journal of Moral Philosophy*, *4*(3), 330–348.

Ridge, M. (2014). *Impassioned belief*. Oxford: Oxford University Press.

Riegel, D. L. (2004). Effects on boy-attracted pedosexual males of viewing boy erotica. *Archives of Sexual Behavior*, *33*(4), 321–323.

Roberts, B. (2009). Disgusting baby Shaker iPhone game. *GoMo News*, 15 May 2009. In G. Young. (2014). *Ethics in the virtual world: The morality and psychology of gaming*. London: Routledge.

Robertson, A. (2019). Virginia's "revenge porn" laws now officially cover deepfakes. *The Verge*, 1 July. Retrieved 10 October 2019 from https://www.theverge.com/2019/7/1/20677800/virginia-revenge-porn-deepfakes-nonconsensual-photos-videos-ban-goes-into-effect.

Robins, K., & Webster, F. (1986). The media, the military and censorship. *Screen*, *27*(2), 57–63.

Rogers, A. (2009). Protecting children on the internet: Mission impossible? *Baylor Law Review*, *61*(2), 323–356.

Ross, W. D. (1930). *The right and the good*. New York: Oxford university Press.

Russell, G. (2008). Pedophilies in wonderland: Censoring the sinful in cyberspace. *Journal of Criminal Law and Criminology*, *98*(4), 1467–1500.

Ryan, M.-L. (1991). *Possible worlds, artificial intelligence, and narrative theory*. Bloomington, IN: Indiana University Press.

Sabur, R. (2018). Harper Lee and Mark Twain banned by Minnesota schools over racial slurs. *The Telegraph*, 12 February. Retrieved 13 December 2019 from https://www.telegraph.co.uk/news/2018/02/12/harper-lee-mark-twain-banned-minnesota-schools/.

Sanson, D. (2016). Philosophy of fiction. *Res Philosophica*, *93*(2), 357–376.

Sapolsky, B. S., Molitor, F., & Luque, S. (2003). Sex and violence in slasher films: Re-examining the assumptions. *Journalism & Mass Communication Quarterly*, *80*(1), 28–38.

Saunders, B. (2016). Reformulating Mill's harm principle. *Mind*, *125*(500), 1005–1032.

Saunders, L. F. (2015). What is moral reasoning? *Philosophical Psychology*, *28*(1), 1–20.

Savage, J. (2004). Does viewing violent media really cause criminal violence? A methodological review. *Aggression and Violent Behavior*, *10*, 99–128.

Schein, C., & Gray, K. (2018). The theory of dyadic morality: Reinventing moral judgment by redefining harm. *Personality and Social Psychology Review*, *22*(1), 32–70.

Schiffer, S. (2003). *The things we mean*. Oxford: Oxford University Press.

Schneider, S. J. (2001). Murder as art/the art of murder. In A. Black (Ed.), *Necronomicon: The journal of horror and erotic cinema, book 4* (pp. 65–85). London: Noir Press.

Schulzke, M. (2011). Reflective play and morality: Video games as thought experiments. In K. Poels & S. Malliet (Eds.), *Vice city virtue: Moral issues in digital game play* (pp. 51–68). Leuven: Acco Academic.

Searle, J. (1979). The logical status of fictional discourse. In J. Searle (Ed.), *Expression and meaning* (pp. 58–75). Cambridge: Cambridge University Press.

Searle, J. (1983). *Intentionality: An essay in the philosophy of mind*. Cambridge: Cambridge University Press.

Searle, J. (1995). *The construction of social reality*. Cambridge, MA: MIT Press.

Searle, J. R. (1996). *Speech acts: An essay in the philosophy of language*. Cambridge: Cambridge University Press.

Seto, M. C., & Eke, A. W. (2005). The criminal histories and later offending of child pornography offenders. *Sexual Abuse*, *17*(2), 201–210.

Seuntjens, W. (2013). The masturbatory fantasy paradox: An overlooked phenomenon? *Journal of Unsolved Questions*, *3*(1), 9–12.

Sheehan, P. (2013). *Modernism and the aesthetics of violence*. Cambridge: Cambridge University Press.

Shelley, J. (1998). Hume and the nature of taste. *The Journal of Aesthetics and Art Criticism*, *56*(1), 29–37.

Shelley, J. (2013). Hume and the joint verdict of true judges. *The Journal of Aesthetics and Art Criticism*, *71*(2), 145–153.

Shiffrin, S. (2012). Harm and its moral significance. *Legal Theory*, *18*, 357–398.

Shimp, T. A., & Stuart, E. W. (2004). The role of disgust as an emotional mediator of advertising effects. *Journal of Advertising*, *33*, 43–53.

Shoard, C. (2011). Human centipede 2 director criticises BBFC over rejection. *Guardian.co.uk*, 7 June. Retrieved 7 December 2019 from http://www.guardian.co.uk/film/2011/jun/07/human-centipede-2-ban-tom-six-spoilers.

Shue, H. (1978). Torture. *Philosophy and Public Affairs*, *7*(2), 124–143.

Shusterman, R. (1984). Aesthetic censorship: Censoring art for art's sake. *The Journal of Aesthetics and Art Criticism*, *43*(2), 171–180.

Sicart, M. (2009). *The ethics of computer games*. Cambridge, MA: MIT Press.

Sidgwick, H. (1981). *The methods of ethics*. Cambridge: Hackett Publishing Company. (Originally published 1874).

Sierz, A. (2001). "The element that most outrages": Morality, censorship and Sarah Kane's blasted. *European Studies: A Journal of European Culture, History and Politics*, *15*, 225–239.

Silk, A. (2014). Why "ought" detaches: Or why you ought to get with my friends (if you want to be my lover). *Philosophers' Imprint*, *14*(7), 1–16.

Simpson, B. (2009). Controlling fantasy in cyberspace: Cartoons, imagination and child pornography. *Information & Communications Technology Law*, *18*(3), 255–271.

Singer, P. (2007). Video crime peril vs. virtual pedophilia. *The Japanese Times*, 22 July. Retrieved 20 July 2019 from https://www.japantimes.co.jp/opinion/2007/07/22/commentary/world-commentary/video-crime-peril-vs-virtual-pedophilia/#.XTWmQZMzau4.

Sinnott-Armstrong, W. (1987). Moral realism and moral dilemmas. *Journal of Philosophy*, *84*(5), 263–276.

Sjöström, A., Sowka, A., Gollwitzer, M., Klimmt, C., & Rothmund, T. (2013). Exploring audience judgments of social science in media discourse: The case of the violent video games debate. *Journal of Media Psychology*, *25*(1), 27–38.

Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. New York: Appleton-Century.

Slotkin, R. (2005). Fiction for the purposes of history. *Rethinking History*, *9*(2/3), 221–236.

Smilansky, S. (1996). The ethical dangers of ethical sensitivity. *Journal of Applied Philosophy*, *13*(1), 13–20.

Smuts, A. (2007). The paradox of painful art. *Journal of Aesthetic Education*, *41*(3), 59–76.

Smuts, A. (2013). The ethics of singing along: The case of "mind of a lunatic." *The Journal of Aesthetics and Art Criticism*, *17*(1), 121–129.

Smuts, A. (2014). Painful art and the levels of well-being. In J. Levinson (Ed.) *Suffering art gladly: The paradox of negative emotion in art* (pp.123–152). Basingstoke: Palgrave Macmillan.

Smuts, A. (2016). The ethics of imagination and fantasy. In A. Kind (Ed.), *The Routledge handbook of philosophy of imagination* (pp. 380–391). London: Routledge.

Sokol-Hessner, L., Folcarelli, P. H., & Sands, K. E. F. (2015). Emotional harm from disrespect: The neglected preventable harm. *BMJ Quality & Safety*, *24*, 550–553.

Stanford, P. K. (2018). The difference between ice cream and Nazis: Moral externalism and the evolution of cooperation. *Behavioral and Brain Sciences*, *41*, e95.

Stern, L. (1965). Fictional characters, places, and events. *Philosophy and Phenomenological Research*, *26*(2), 202–215.

Stern, L. (1967). On make-believe. *Philosophy and Phenomenological Research*, *28*(1), 24–38.

Stock, K. (2005). Resisting imaginative resistance. *Philosophical Quarterly*, *55*(221), 607–624.

Stock, K. (2016). Imagination and fiction. In A. Kind (Ed.), *The Routledge handbook of philosophy of imagination* (pp. 204–216). London: Routledge.

Stone, G. R. (2007). Sex, violence, and the first amendment. *The University of Chicago Law Review*, *74*, 1857–1871.

Stoner, I. (2020). Barbarous spectacle and general massacre: A defence of gory fictions. *Journal of Applied Philosophy*, *37*(4), 511–527.

Strasburger, V. C., & Wilson, B. J. (2014). Television violence: Sixty years of research. In D. A. Gentile (Ed.), *Media violence and children* (2nd ed., pp. 135–177). Santa Barbara, CL: Praeger.

Stueber, K. R. (2011). Imagination, empathy, and moral deliberation: The case of imaginative resistance. *The Southern Journal of Philosophy*, *49*(Spindel Supplement), 156–180.

Sussman, D. (2005). What's wrong with torture? *Philosophy & Public Affairs*, *33*(1), 1–33.

Symonds, G. (2008). *The aesthetics of violence in contemporary media*. New York: Continuum Int.

Tan, A. S., & Scruggs, K. J. (1980). Does exposure to comic book violence lead to aggression in children? *Journalism and Mass Communication Quarterly*, *57*(4), 579–583.

Tate, T. (1990). Child pornography: An investigation. London: Methuen.

Tate, T. (1992). The child pornography industry: International trade in child sexual abuse. In C. Itzin (Ed.), *Pornography: Women, violence and civil liberties* (pp. 203–216). Oxford: Oxford University Press.

Tavinor, G. (2005). Videogames and interactive fiction. *Philosophy and Literature*, *29*(1), 24–40.

Thomasson, A. L. (1999). *Fiction and metaphysics*. Cambridge: Cambridge University Press.

Thompson, K. M. (2004). Violence, sex, and profanity in films: Correlation of movie ratings with content. *Medscape General Medicine*, *6*(3), 3.

Thomson, J. J. (2011). More on the metaphysics of harm. *Philosophy and Phenomenological Research*, *82*(2), 436–458.

Todd, C. S. (2009). Imaginability, morality, and fictional truth: Dissolving the puzzle of “imaginative resistance.” *Philosophical Studies*, *143*, 187–211.

van Inwagen, P. (1977). Creatures of fiction. *American Philosophical Quarterly*, *14*, 299–308.

van Inwagen, P. (1983). Fiction and metaphysics. *Philosophy and Literature*, *7*(1), 67–77.

van Ooijen, E. (2011). Cinematic shots and cuts: On the ethics and semiotics of real violence in film fiction. *Journal of Aesthetics and Culture*, *3*(1), 1–16.

Vecsey, Z. (2015). A representational account of fictional characters. *Analysis and Metaphysics*, *14*, 68–79.

Vincent, J. (2019). New AI deepfake app creates nude images of women in seconds. *The Verge*, 27 June. Retrieved 10 October 2019 from https://www.theverge.com/2019/6/27/18760896/deepfake-nude-ai-app-women-deepnude-non-consensual-pornography.

Vingehoets, J. J. M., Bylsma, L. M., & de Vlam, C. (2013). Swearing: A biopsychosocial perspective. *Psychology Topics*, *2*, 287–304.

von Tunzelmann, A. (2013). Zero dark thirty’s torture scenes are controversial and historically dubious. *The Guardian*, 26 January. Retrieved 25 November 2019 from https://www.theguardian.com/film/filmblog/2013/jan/25/zero-dark-thirty-reel-history.

Waddington, D. I. (2007). Locating the wrongness in ultra-violent video games. *Ethics and Information Technology*, *9*(2), 121–128.

Waldron, J. (2012). *Dignity, rank, and rights*. Oxford: Oxford University Press.

Walton, K. (1990). *Mimesis as make-believe: On the foundations of the representational arts*. Cambridge, MA: Harvard University Press.

Walton, K. (1994). Morals in fiction and fictional morality (part 1). *Proceedings of the Aristotelian Society*, *Supplementary Volumes*, *68*, 27–51.

Walton, K. (2006). On the (so-called) puzzle of imaginative resistance. In S. Nichols (Ed.), *The architecture of imagination: New essays on pretense, possibility, and fiction* (pp. 137–148). Oxford: Oxford University Press.

Warburton, W. A., Roberts, D. F., & Christenson, P. G. (2014). The effects of violent and antisocial music on adolescents. In D. A. Gentile (Ed.), *Medial violence and children: A complete guide for parents and professionals* (2nd ed., pp. 301–328). Santa Barbara: Praeger.

Ward, D. V. (1990). Philosophical issues in censorship and intellectual freedom. *Library Trends*, *39*(1/2), 83–91.

Waters, J. (2005). *Shock value: A tasteful book about bad taste*. Philadelphia: Running Press Adult.

Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychology Review*, *20*(2), 158–177.

Weatherson, B. (2004). Morality, fiction and possibility. *Philosophers’ Imprint*, *4*(3), 1–27.

Weinberg, J. M., & Meskin, A. (2006). Puzzling over the imagination: Philosophical problems, architectural solutions. In S. Nichols (Ed.), *The architecture of imagination: New essays on pretense, possibility, and fiction* (pp. 175–202). Oxford: Oxford University Press.

Weisberg, D. S. (2016). How fictional worlds are created. *Philosophy Compass*, *11*(8), 462–470.

Wheater, M. (1993). Mozart's health and final illness. *Dalhousie Review*, *73*(2), 213–223.

White, A. (2006). *Virtually obscene: The case for an uncensored Internet*. London: McFarland & Company, Inc.

White, H. (2005). Introduction: Historical fiction, fictional history, and historical reality. *Rethinking History*, *9*(2/3), 147–157.

Williams, C. (2007). Some questions on Hume's aesthetics. *Philosophical Compass*, *2*(3), 157–169.

Williams, Z. (2003). Panic on the screens. *The Guardian*, 14 January. Retrieved 7 December 2019 from http://www.guardian.co.uk/media/2003/jan/14/newmedia.childrensservices.

Woodcock, S. (2015). Comic immoralism and relatively funny jokes. *Journal of Applied Philosophy*, *32*(2), 203–216.

Yablo, S. (2002). Coulda, woulda, shoulda. In T. S. Gendler & J. Hawthorne (Eds.), *Conceivability and possibility* (pp. 441–492). Oxford: Oxford University Press.

Yagisawa, T. (2001). Against creationism in fiction. *Philosophical Perspectives*, *15*, 153–172.

Young, A. (2000). Aesthetic vertigo and the jurisprudence of disgust. *Law and Critique*, *11*(3), 241–265.

Young, G. (2013). Enacting taboos as a means to an end; but what end? On the morality of motivations for child murder and paedophilia within gamespace. *Ethics and Information Technology*, *15*(1), 13–23.

Young, G. (2014a). *Ethics in the virtual world: The morality and psychology of gaming*. London: Routledge.

Young, G. (2014b). A meta-ethical approach to single-player gamespace: Introducing constructive ecumenical expressivism as a means of explaining why moral consensus is not forthcoming. *Ethics and Information Technology*, *16*(2), 91–102.

Young, G. (2015). Violent video games and morality: A meta-ethical approach. *Ethics and Information Technology*, *17*(4), 311–321.

Young, G. (2017). Integrating poor taste into the ongoing debate on the morality of violent video games. *The Computer Games Journal*, *6*(4), 227–237.

Young, G. (2019). An expressivist account of the difference between poor taste and immorality. *Ethical Theory and Moral Practice*, *22*(2), 465–482.

Young, G., & Whitty, M. T. (2011). Progressive embodiment within cyberspace: Considering the psychological impact of the supermorphic persona. *Philosophical Psychology*, *24*(4), 537–560.

Young, G., & Whitty, M. T. (2012). *Transcending taboos: A moral and psychological examination of cyberspace*. London: Routledge.

Zegers, R. H. C., Weigl, A., & Steptoe, A. (2009). The death of Wolfgang Amadeus Mozart: An epidemiological perspective. *Annals of Internal Medicine*, *151*(4), 274–278.

Zvolenszky, Z. (2013). Abstract artifact theory about fictional characters defended—Why Sainsbury's category-mistake objection is mistaken. *Proceedings of the European Society for Aesthetics*, *5*, 597–612.

Index

Note: Page numbers in italics refers to main area of reference

abbreviated descriptor, *31–32*, 34, 162
abstract artefact theory. *See* creationism (fiction)
abstract entities, *23*, 26, 28, 162, 164
accidental truth (of fiction), 30, 160, *162–65*
ACRB. *See* Australian Classification Review Board
aesthetic appreciation, 11, 14, 68, 72, 74, 156, 227, 230
aestheticization of violence, 12, *72–73*
alethic commitment, *3–4*, 198
Ali, R., 120, 127–28
Amadeus, 15, 159–61, 163–65, 169, 171–72, 175n3, 229; Mozart, 15, 159–65, 169, 171, 174n1, 175n4; Salieri, 15, 159–65, 169, 171–73, 174n1, 175nn3–5, 229
Anderson, C.A., 106, 108, 110, 113n9
anti-semitic/semitism, 171, 173
Aristotle, critique of fiction, 100–102, 113n8
Australian Classification Review Board, *5*, 16n2, 238
authoritative breakdown, 137
authorized game world, *126–28*, 131n5, 134, 138, 161
Bartel, C. *40–42*, 46, 115; and Cremaldi, A., 42, 63, 89, 192–93, 195, 197–98
BBFC. *See* British Board of Film Classification
Bill Sikes. *See* literature, Oliver Twist
books. *See* literature
Booth, W.C., 55, 59, 75n12, 139, 168
British Board of Film Classification, 5, 16n2, 238
Buscemi-Lawrence deepfake, *183–84*, 229
Bushman, B.J., 105–6, 109, 113n9

Carroll, N., 6, 69, 133
categorical imperative, 183, 186n8, 236. *See also* hypothetical imperative
catharsis / cathartic, *100–102*, 106, 113n8
CEE. *See* constructive ecumenical expressivism
censorship, 6–7, 13, 16, 100, 106, 231; definition, 231–32, 239n8; freedom, 233, 236–37; harm, 232; truth, 236
Charles Dickens. *See* literature, Oliver Twist
Cherry, C., 188–90, 193–94
child pornography, 8, *39–42*, 44–46, 177, 195; as child abuse, 5–6, 35,

38, *42–45*, 52n5, 112, 196–97; definition, 52; law, 52, 196. *See also* virtual child pornography; virtual pedophilia

Child Pornography Protection Act (CPPA). *See* legislation

codes of appropriateness, 150

coincidently true. *See* accidental truth (of fiction)

comic amoralism, 6

concreta, 23, 25–26

conditions for existence, universal and relative, 21–23, 26, 33n2, 34n4

conditions for harm. *See* harm conditions

consent, 52n7, 66, 94–95, 97n12, 136–38, *178–82*, 185, 186n11, *199–203*, 205, 227

constructive ecumenical expressivism (or CEE), 7, 13, 16, *210–20*, 222, 223n2, 224–25, 228–29, *233–38*; conflicting norms, 213; constructed social (or moral) norm, 7, 24, *212–13*, 228–29, 235; moral error, 213; robustness, 13, 15, *212–16*, 224n9, 234–36, 238

contradiction, problem of, *208–9*, 222, 233

Cooke, B., 3, 63, 168, 187, 195, 197–98

co-opted (imaginative resistance as moral marker), 131, *135*, 138

corporeal / corporeality, 22–23, 32, 37, 162, 164

counterfactual, 160–63, 171, 173–74, 188, 192, 194, 203, 206n2

counterfactual-comparison approach (to harm), *80–86*, 89, 90–91

creationism (fiction), 24–26

cultural harm, 112, 231

Currie, G., 29–30, 34n10, 69

de dicto attitude, 211–12, 214–15, 218–20, 224n9, 234. *See also* de re attitude

deepfake (pornography): disclosure, 180–83, 185, 202; disrespect, 15, 183–85, 186nn9, 11; harm, 182; nonconsensual pornography, 15, 178–82, 184–85, 186n11, 229, 238n3

deepfake (private); disrespect, 200–201; harm, 200; nonconsensual pornography, 200

depiction dependent, 118

de re attitude, 211–12, 214–15, 217, 219–20, 234–35, 238. *See also* de dicto attitude

descriptivism, 31–32

desensitization. *See* harm conditions, $\text{SHC}_{\text{desensitization}}$

desideratum (fantasy), 194

Dickens, Charles, 1, 7, 19–20, 23, 28–29, 37, 66, 133

Di Muzio, G., 37, 52n1, *101–2*, 117

directive speech act, 152–53

disapprobation, 6

disesteem, 205, 224n8, 230

doctrine, 234

dogma, 234–35, 237

dramatization (of historical events), 163–66, 169, 171

Driver, J., 153–54, 205

Durkheim, E., 222

to elicit (arousal / enjoyment), 13, 42, 100, *121–27*, 133–34, 155, 191, 196, 206n6, 227

embellishing, fantasy. *See* fantasy, embellishing

enjoyment, differentiating: $\text{enjoyment}_{\text{(sim)}}$, *125–31*, 227, 231; $\text{enjoyment}_{\text{(sub)}}$, *125–31*, 216, 227–28, 237; $\text{M}_{\text{(enjoyment)}}$, 124–25; $\text{M}_{\text{(substitution)}}$, 125; strong sense, 116, 119, 124–25; weak sense, 122, 131n3

entailment (logical / conceptual), 37–38, 117, 141, 195, 216, 226

epistemic relationship, 61, 144, 146n11, 164–65, 226

etiquette, 154

existence, universal and relative conditions. *See* conditions for existence

fantasy: active, 193–94; disrespect, 191, 203–5, 216, 230; embellishing, 193–94, 204, 230; harm, 202; idle, 16, 193–99, 201–5, 216–17, 230; involuntary, 192–93; as non-consensual pornography, 201; passive, 193–94; self-gratifying, 188–89, 194; self-referential, 189–90; spontaneous, 193; surrogate, 191, 193–94, 197, 199, 202–5, 217, 229–30; usurping, 188–90; voluntary, 192–94
Feinberg, J., 80–81, 111, 113n3, 175n15, 182
Ferguson, C.J., 12, 107–9, 113n9, 158n7
fictional truth, 126–27, 129, 137, 139
fictitious video games: Child Sexual Assault, 10; Luck's jewel thief, 45
fictive intent, 29, 162
fictively imagine, 3–4, 55, 127, 130, 134–35, 137–38, 140–41, 144, 161–62, 165, 194, 198, 225, 227
fidelity constraint, 27–28, 33–34, 160–61, 163–64, 166, 169
films: Blade Runner, 50; Bonnie and Clyde, 166; Braveheart, 167–68; Cannibal Holocaust, 49; A Clockwork Orange, 12, 72; The Good, the Bad and the Ugly, 73; Human Centipede 2, 5, 69; Kill Bill (1 and 2), 12, 72; Passion of the Christ, 73; Pink Flamingos, 155; Pulp Fiction, 48–49; A Serbian Film, 5, 10–11, 35–36, 51, 117; Straw Dogs, 70, 117, 125–27; Triumph of the will, 63; The Wild Bunch, 70; Zero Dark Thirty, 168
fixed norms, 55–56, 73, 139. *See also* nonce beliefs
freedom: exercise of, 233, 237; of expression, 13, 39, *233*, 236–38; from / for (on censorship), 236

game world, 119–20, *126–31*, 131n5, 134–42, 227
Gaut, B., 3, 13, 70, 73, 123, 131n2, 199, 239n4
Gendler, T.S., 69, 137, 139, 141–42, 146n6
Gert, B., 79, 84, 96n7, 168
God's law, 151, 211, 218–19, 221
Gunter, B., 107

harm, evidence of: cartoons, 106–7, 111; lack of consensus, 12, 14, 99, 109, 111, 199, 227, 232; media violence (general), 99, *104–8*, 190, 238; media violence and crime, 108–10; novels, 107, 110–11; song lyrics, 109–11; 'violent' video games, 12, 93, 108–9
harm conditions: HC_{causal}, 86–87, 89; HC_{risk}, 90; $HC_{suffering}$, 85, 91; supplementary harm conditions: $SHC_{behavior}$, 86–87, 90, 104; SHC_{belief}, 86–90, 104; $SHC_{conation}$, 87–90, 104; $SHC_{desensitization}$, 88, 94
harming others, *90–95*, 108, 227
harm through omission, 86–88
hermeneutic recalibration, 69, 75n10, 134
Hershfield, J., 189–91, 193, 195
Hindley, Myra, 156, 218
Holmes, Sherlock, 20–21, 24–26, 29, 33, 34n9, 134
Holocaust, 17n7, 37, 167, 170; art, 11
homosexual, 17n4, 68, 136–38, 214
horror, 69, 101, 133, 152–53; film, slasher, 12, 72, 75n8, 96n9, 102
Huesmann, L.R., 7, 106, 109, 113n9
hypothetical imperative, 158n5. *See also* categorical imperative

image-based sexual abuse, 15, 178, 182
imaginative / fictional experience, 129, 133
immorality conditions: IC_1, 77–78, 99, 111; IC_2, 92–95, 99, 111
immoral world view. *See* IWV; $IWV_{revised}$
inalienable rights, 233, 237

incorrigible social meaning, 65–66, 68, 75n9, 227, 231
intersubjective (norm or standard), *212–14*, 220, 228–29, 234–35
intuition, 9–10, 19, 23–24, 117–18, 124, 129, 145
IWV, 9, 14, 57–58, *60–65*, 68, 72, 74, 140, 147, 156, 173, 211, 215–16, 232–37
IWW$_{\text{revised}}$, *64–67*, 74, 75nn5–6, 140, 147

Kajtár, L., 20, 27, 163–64
Kant, I., 94, 158n5, 183, 186n8, 203, 205, 236
Kelly, Ned, 165–66
Kershnar, S., 88, 131n1, 195, 197–99
Kieran, M., 69, 103
king, 31

legislation: Child Pornography Prevention Act (CPPA), 39–40; Miller test (for obscenity), 112n3; Obscene Publications Act, 103, 113n5; PROTECT Act, 39–40, 52nn6–7
Levi, Primo, 167, 170, 175n9
Lindbergh, Charles, 169, 171–73, 175n12
literature: The Adventures of Huckleberry Finn, 47, 50; American Psycho, 36, 70, 133; Beowulf, 24; Catch-22, 55–56, 69; Fatherland, 160–61; The Jungle Book, 21; Lolita, 46, 52, 57–58; Oliver Twist, 1, 7, 19–20, 28; The Plot Against America, 169, 171, 173; Pride and Prejudice, 56; The Prisoner of Azkaban, 21; The Sign of Four, 24; A Study in Scarlet, 24–25; To Kill a Mockingbird, 47, 50
Luck, M., 45–46, 53n9

MacManister, Kitty Jane, 29–31
make-believe/belief, 15, 27, 55, 59, 69, 127, 129, 131n5, 134, 136–40, 142, 144, 146nn9, 11, 164–65, 168, 174, 227, 235
manga, 12, 107, 111
manipulation, 15, 139, 146n9, 177–84, 186, 200
Marley, Jacob, 66, 136
materially innocent. *See* morally non-innocent
McCormick, M., 91–93, 140, 232–33
media violence: definition, 105; exaggerates real violence, 106
Mill, J.S. 79, 85–86, 94; on censorship, 234–36
Miller test (for obscenity). *See* legislation
Mind of a Lunatic (song), 129–30
mind-to-world direction of fit. *See* world-to-mind direction of fit
misrepresent: history, 170–71; a morally pertinent matter, 221, 233; sexual activity, 179–81, 183, 186, 201–3
molestation, 128, 196
moral: appraisal, 77, 115; attitude, 143, 212–13, 224, 234, 238; corruption, 10, 102, 104, 232; deviance, 137, 139; disapproval, 6, 102, 143, 157, 210, 217–18, 220; disorientation, 140; edification, 14, 57, 67–69, 74, 75n12, 230; insensitivity, 14, 67–68, 71, 74–75, 127, 147–48, 154, 172, 174, 227, 230–31. *See also* WV$_{\text{insensitive}}$; property, 141–42, 146, 213. *See also* nonmoral property; realism, 210, 214; relativism, 214, 234–36; wisdom, 131
moral-fiction, 138–40
morally non-innocent, 111, 231
morphed image, 39
movies. *See* films
Mozart. *See* Amadeus, Mozart
murder: actual, 120–21, 123, 125, 227; virtual, 120–22, 124

Nabokov, V., 46, 58. *See also* literature, Lolita

Nancy (Oliver Twist), 1, 7, 20, 23, 28, 37, 51, 133
Napoleon, 28
Nazi, 17n7, 63, 122, 167, 171, 173
necessary anachronism
nonce beliefs, 55–56, 139. *See also* fixed norms
noncomparison approach (to harm), 80, 82–84
nonmoral property, 141, 143–44. *See also* moral, property
nonpredication theory, 19, 136
non-pro tanto harm, 227–28. *See also* pro tanto harm
normative biological state, 82–86, 90–91
"no smoke without fire" objection, 172

objectified (norm), 7, 156, 213, 235
obligation / obliged, 15, 77, 78, *152–54*, 158, 172, 184–85, 190, 205, 221, 230, 233
obscene. *See* obscenity
Obscene Publications Act. *See* legislation
obscenity, 11, 40, 99, *102–4*, 112nn2–3, 113n6, 232
offense / offensive / offensiveness, 5, 11, 48–50, 65, 75n7, 103–4, 110, 112; and poor taste, 148–51, 156, 158n3, 207, 217, 219, 222, 223n3, 224n10, 231
ontological equivalence, 42–43
ontology / ontological, 20–21, 23, 28, 41, 43–44, 60–61, 75n4, 119, 123, 179, 196, 214, 226
ought, types of, 152–53, 158nn5–6, *208–9*, 221–23, 225, 236, 238

Palahniuk, Chuck, 38, 157
parasitic (on morally pertinent), 16, 170, 220
Patridge, S., 10, *65–67*, 108, 129, 183, 227
Peckinpah, Sam, 70, 117
pedophilia. *See* child pornography; virtual child pornography; virtual pedophilia
"Piss Christ," 11, 17n7
Plato, critique of fiction, 100, 112n1
playing / playful, 11, 66, 92, 105, 109, 117, 120–21, 125, 130, 207
poor taste: good bad taste / bad bad taste, 155–56, 231
"pop-out" (imaginative resistance), 137
pornography: definition, 41; ontological status, 41
precautionary principle, 77, 93, 227
pretense, 29, 197, 199
principle of minimal departure, 56
procure (enjoyment), 115–16, 126, 128–30, 226
prohibition, 6, 39, 152, 189, *208–9*, 222, *231–33*
pro tanto harm, 79, 85, 94, 96, 156, 227–28. *See also* non-pro tanto harm
PROTECT Act. *See* legislation
prurient interest, 112n3
public expression (of disrespect), 204–5, 216, 230
pure fiction, 19–20, *27–33*

quasi-realism, 214
Queen: Elizabeth I (Elizabethan), 35, 165; Victoria (Victorian), 102, 192

Rabbit, Jessica, 191, 193, 198
racial slurs, 14, 36, 38, *47–51*, 226
RapeLay. *See* video games
Rea, M., 41, 179
reasonable interpretation, 14, 57, 60–62, 64, 67, 71, 74, 75, 140, 147, 173, 226, 230
recommendation and requirement (differentiation), 152–53, 222
reliable indicator, 135, 145, 228
reputation, 15, 172–74, 175n15, 182–83, 200, 202, 229, 238n3
res cogitans, 21–22
revenge porn, 178, 180, 185n1
ring true test, 165

risk of harm. *See* harm conditions, HC_{risk}
robust relativism, 234–36
Rowling, J.K., 21, 28, 61

sadomasochism, 123–24
Salieri. *See* Amadeus, Salieri
Searle, J., 25, 31, 152, 191, 223
secular duty, 151, 211, 218–19
Serpico, Frank, 26, *29*, 33
sexual activity, 39–40, 43, *178–82*, 185, 186n5
sexual voyeurism, 195–96
Simpson, Marge, 2, 187–88, 191, 195, 198
simulate / simulation, 17n4, 78, *123–25*, 191
singing, 129–30
skin job, 50
slasher films. *See* horror, film, slasher
slurs. *See* racial slurs
Smuts, A., 129, 133, 193, 205, 224n8
solicitation (child), 128, 195–96
speech act theory, 25–26
Star Trek, 59–60, 165; Kirk, 135, 136–38, 145n4; USS Enterprise, 135–36, 138
Star Wars, 127, 137; Skywalker, 56, 126; Vader, 56
suberogatory, 15–16, 68, 147–48, 153–55, 157, 186n9, 205, 218, 222–23, 225–26, 230–32
supervenience / supervenient, 141–44, 146; strong, 142; weak, 142, 144
supplementary harm conditions. *See* harm conditions, supplementary harm conditions

taboo(s), 70, 103, *122–26*
Tallahatchie bridge, 29–30
Tarantino, Quentin, 12, 48, 72
taste. *See* offense / offensive / offensiveness, and poor taste; poor taste
temporal-comparison approach (to harm), 80–86, 89–91
tolerance: of censorship, 235; personal, 15, 135, 145, 228
Tolstoy (and War and Peace), 28
transparent (deepfake_p), 181
treatment of morally pertinent matter, 15–16, 148, 150, *154–57*, 158n7, 203, 207, 210, *219–23*, 231
trivializing (the morally pertinent), 62–63, 148–53, 217–21, 231
true of / about, 32–33, 34n12, 130, 136–37, 143–44, 162, 164
truth-aptness, 20, 136, *144*, 146, 162, 164, 211

unauthorized game world. *See* authorized game world
understanding, of fictional-x_i, 119–22, 124
universality, 21, 33n2, 226

verisimilitude, 170–71, 173, 234–35
vicarious / vicariously, 14, 103, 115, *125*, 127–28, 194, 197, 227
video games: Baby Shaker, 6; Call of Duty, 6; Enzai: Falsely Accused, 6, 117; Ethnic Cleansing, 9, 58, 60, 64; Grand Theft Auto, 9, 12, 62; Hatred, 35; Heavy Rain, 12; Jesus Strikes back, 227; Mafia Wars, 65; Manhunt 2, 12, 35; Playing History 2: Slave Trade, 151; Postal 2, 6; Silent Hill, 45
virtual child pornography, 39–40, 43–45
virtual pedophilia, 39–46; non-pornographic, 44–46. *See also* child pornography; virtual child pornography

Waite, Terry, 189
Walton, K., 16n1, 56, 119–20, 126, 131, 134, 136, 138, 140, 142, 161
Waters, J., 155–56, 231
well-being (psychological), 104
"why would you even think that?" objection, 172–73
work world. *See* game world
world-to-mind direction of fit, 192, 198; $WV_{insensitive}$, *67–68*, 72–74, 140, 147, 173–74. *See also* moral, insensitivity

About the Author

Garry Young is a senior lecturer in philosophy in the School of Historical and Philosophical Studies, University of Melbourne. His research interests include examining the morality of enacting violent or otherwise taboo activity within video games, embodied cognition, and delusional beliefs. He is the author of *Resolving the Gamer's Dilemma* (2016, Palgrave Macmillan), *Ethics in the Virtual World* (2014, Routledge), and *Transcending Taboos: A Moral and Psychological Examination of Cyberspace* (2012, Routledge; coauthor Monica Whitty).